Journal of
Experimental Psychology

VOL. 32, NO. 2 FEBRUARY, 1943

THE EFFECT OF DELAYED REINFORCEMENT UPON THE DIFFERENTIATION OF BAR RESPONSES IN WHITE RATS[1]

BY C. THEODORE PERIN*

Yale University

In a previous experiment (**8**) the writer has found evidence which strongly suggests that food reinforcement must follow in twenty seconds or less a bar-pressing response made by a hungry rat in order for the response to be learned by the animal. Unfortunately, the only measure of habit strength available in that experiment was the response latency, the elapsed time after the animal was given access to the bar until it pushed the bar.

It is recognized that the generality of the conclusions would be greater if several measures of habit strength should be found to yield similar results in terms of a gradient of delayed reinforcement, and the present experiment was planned and carried out in order to secure results from more than one measure of habit strength. In the first delayed-reinforcement experiment by the writer (**8**) latency was the only measure available, because only one response, a downward push on a horizontal bar, was recorded in the situation. If additional responses, incompatible with the correct response, were made available to the animal and provision were made for recording these responses, an additional measure of habit strength would be obtained, namely, the percent of correct responses occurring in relation to the total number of responses made.

It was thought that the degree of extinction and inhibition could be lessened if at first the two responses were trained, either response

[1] This investigation is a part of the coördinated research program of the Institute of Human Relations, Yale University. The writer is deeply indebted to Professor Clark L. Hull for generous advice and aid. The present paper is a portion of a dissertation presented to the Faculty of the Graduate School of Yale University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

* From the Department of Psychology, Institute of Human Relations, Yale University.

being reinforced with food immediately, and later they were differentiated with different degrees of delayed reinforcement. Presumably the gradient in the first experiment was shortened by the presence of the inhibiting factors, and, if so, the partial elimination of these factors should result in learning with more protracted periods of delay than were found to be effective in the first experiment. The line of reasoning underlying this prediction is substantially as follows: If at the beginning of learning the strength of the response tendency were at or near a zero value as in the first experiment, the small increments of habit strength resulting from reinforcement after a relatively long delay period (a half minute for example) would accumulate very slowly. Now if every response, in addition to the factor of reinforcement, also involves increments of inhibition of the bar-pressing response as well as increments of extinction to the response of crossing the box to the food tray, it is quite possible that the smaller increments of habit strength would be completely overridden by the stronger inhibiting tendencies, and thus the small accumulation of primary habit strength would never be reflected in the final measure of response occurrence.[2] By permitting the response to acquire a degree of strength and then later requiring that it be differentiated from a second response of equal or greater habit strength under conditions of delayed reinforcement, the inhibiting effects resulting from the delay period would be less, presumably, and thus a more accurate picture of the reinforcement process would result.

Subjects and Apparatus

The final results represent the data obtained from a total of 150 albino rats approximately 90–120 days old at the beginning of training. These animals comprised six experimental groups, 25 rats in each group, corresponding to six different delay-of-reinforcement intervals which were utilized (*viz.*, 0, 2, 5, 10, 20, and 30 sec.). Additional animals were discarded early in the procedure because of extreme timidity or because of failure to learn the initially trained responses.[3]

The standard diet for these animals consisted of Purina Dog Chow biscuits, which also constituted the main portion of the reward pellets given in the training situation.

The apparatus included a new modification of the sound-proofed experimental boxes which had been used by Ellson (2), Williams (10), and Perin (6), extensive changes having been made to meet the requirements of the present experiment.

The heavy wooden outer shell of the box,[4] the layer of sound-proofing material and the intermediate chamber within this sound-proofing layer were the same as had been used in previous experiments. The inner compartment in which the rat was placed was made even smaller than that used in the earlier study by the writer (8) and was about one-fifth the size of the experimental

[2] This line of reasoning is similar to the 'fatigue hypothesis' advanced previously by Mowrer (5) and by Miller and Dollard (4).

[3] The criteria for elimination of such animals were the same as those used by the writer in an earlier study (8). No animals were discarded after they had been assigned to the different experimental groups late in the procedure.

[4] Although the following description mentions only one experimental box and apparatus accessories sufficient for running animals one at a time, two identical boxes were used which economized on time by permitting two animals to be run concurrently.

compartment originally used in the box. This small inner chamber was 6 in. wide, 5 in. high and about 9 in. long. Reducing the compartment to this size served to decrease the variability of an animal's responses by restricting movement to a small area. A vertically sliding metal door 6¼ in. inside this chamber could be raised or lowered by the experimenter by means of a silk fishline running through glass beads to the top of the box. When closed, this door prevented access to the bar and food.

The end of the animal compartment opposite the entrance was covered by a brass panel in which a slot 20 mm wide by 8 mm high was cut equidistant from the sides of the animal compartment and about 45 mm from the floor. The bar which the animal had to learn to press differed from those used in previous experiments in that it was cylindrical in shape, projecting into the compartment through the slot in the brass panel. It was 7.5 mm in diameter with the center drilled out to make a hole into which bits of moist food could be inserted. On the floor, directly beneath the bar, was a food tray into which food reward pellets could be delivered through a small brass tube projecting from the brass panel.

When the bar was in position to be pushed by a rat it projected 20 mm into the box. A pressure of 3 gm on either side of the bar moving it either right or left a distance of 2 mm was necessary to close electrical contacts and thus constitute a response. The bar was mounted on a movable platform sliding in two brass channel grooves. These grooves were carefully machined and the points where the bar carriage came in contact with the channel supports were constantly lubricated with heavy machine oil in order to permit the bar to slide easily and silently. A sponge rubber block served to stop the backward movement of the bar with a minimum of noise.[6] A brass rod running to a door-bell magnet mounted outside the box served to hold the bar in position, inserted into the animal compartment, until the magnet was energized by the occurrence of the correct response. As soon as the catch on the bar was released by the magnet, a small spring attached under tension to the bar withdrew it from the animal compartment, thus effectively preventing any further responses.

Right and left responses were recorded on a constant-speed waxed-paper record by two magnetic signal markers. A third marker could be operated by the experimenter to record the moment at which the starting gate was raised.

A one-RPM Telechron motor with attached sweep arm was used to operate the food-release mechanism which delivered the food pellets after the proper delay interval following the correct response.

At the stage in the procedure where a delay period preceded reinforcement of the bar response, the operation of the apparatus was essentially as follows. One of two knife switches was closed, depending on whether a left or right bar response was "correct" for the given animal, and a selector switch was set corresponding to the reinforcement delay for that animal. The rat was then given access to the bar by the raising of the starting gate. If the first responses were incorrect (*i.e.*, bar responses in the non-rewarded direction) they were registered on the waxed-paper record. When a correct response occurred it was recorded at a different position on the record. Simultaneously with the occurrence of this correct response the magnet controlling the bar retraction was activated and the bar was automatically pulled from the animal compartment. The action of the response-recording marker also served as a relay which in turn closed a 110 volt A.C. relay, the latter serving to operate the Telechron delay timer. The timer continued running through the use of a holding relay until one revolution was completed at which time the current was automatically broken; this set the timer for the next trial.

Simultaneously with the occurrence of a correct response a fourth relay, wired in parallel with the bar release magnet, closed and this in turn operated a 110 volt D.C. holding relay, partially completing the circuit for the delivery of the food pellet. When the sweep on the Telechron timer made contact with a brass screw following the proper delay-of-reinforcement interval, the release mechanism, located on the opposite side of the brass panel from the animal

[6] Extreme care was taken in the design of the apparatus to keep noises at a minimum. It is believed, on the basis of the Bugelski experiment (1) and other observations, that these auditory stimuli are capable of acquiring much strength as secondary reinforcing agents. All relays and electromagnets which were operated by bar responses were located outside the experimental box. Bar retraction was made as silent as possible lest the food-release mechanism noises which themselves become powerful as agents of secondary reinforcement should generalize to these bar noises and permit these latter to function also as secondary reinforcement.

compartment, dropped a pellet from the food magazine into the tray directly in front of the animal, at the same time making a distinct clicking noise.[6]

The food pellets themselves, often described before, were cylindrical in shape, weighing about .05 gm each. They were composed chiefly of ground Purina Dog Chow.

PROCEDURE

Habituation to the Food-Release Mechanism

After a definite feeding rhythm had been established and the rats had become accustomed to eating in the experimental boxes, two days were spent in training them to retrieve the pellets of food which were released into the food tray when the experimenter operated the food-release device. This procedure has been described in detail before (**6**). Sixty pellets were given on each of the two habituation days.

Initial Training on the Bar

As suggested in the introductory paragraphs, a salient feature of the present procedure was the establishment of a considerable degree of habit strength for the bar response before a delay period was introduced. This preliminary bar training was given on the day following the completion of the habituation to the food mechanism.

The bar was first inserted into the experimental compartment and the switches were set so that it did not withdraw when either right or left pushes were made.[7] Additional switches were set so that a food pellet was released into the tray simultaneously with the pressing of the bar in either a right or left direction. The drilled-out center of the bar was filled with a paste made of ground Purina Dog Chow.

After the animal had been placed in the experimental box it would attempt to remove the moist food from the hollow bar. This was sufficient, on the average, to move the bar far enough toward the right or left to cause the electrical contact to be made, resulting in the delivery of a food pellet. As soon as the pellet had been consumed the animal was returned to its living cage. The other three animals being run concurrently with the first were each in turn put into the experimental boxes for this same procedure. In this way each animal was given 50 reinforced bar-pressing trials. After the first few trials the habit became so strong that it was no longer necessary to use the moist food as an additional stimulus to bring about the response. The day following, 10 more trials were given in the manner just described. This was done to restore the habit to about the same strength as it had been at the end of the previous day's training and also to test for spontaneous reversal of direction preference for bar pushing. The number of responses in each direction was determined for the last 20 bar responses of each rat. The direction occurring least often was selected as the 'correct' response for the reversal training under the conditions of delayed reinforcement.

Two factors entered into the placement of animals in the different delay groups. The animals were distributed equally among experimental groups on the basis of latency scores for the last 20 responses under conditions of simultaneous reinforcement, it being considered that latency constitutes one measure of habit strength and that in this case it reflected the learning ability of individual animals. The other basis for matching the groups was the direction of preference for bar pushing, including the percent of occurrence of the preferred response.

---

[6] The writer has recognized previously (**8**) that the click of the food mechanism serves as a valuable 'cue' to the rat that food has been delivered, and reduces to a minimum the time necessary for the animal to procure the food pellet. In order to make this click more audible to the animal the food-release device was moved from its position outside the box and instead was attached directly behind the brass panel inside the experimental box.

[7] Obviously the bar should not have withdrawn following each response at this stage of the procedure. If the bar withdrew immediately after being pressed in either direction and this withdrawal had been accompanied simultaneously with the delivery of a food pellet, it would have been possible for the bar withdrawal to become a stimulus possessing secondary reinforcement characteristics which presumably would have nullified to a considerable extent the effect of the delay period later in the procedure when the two responses (right and left) were to be differentiated.

Differential-Response Learning with Delayed Reinforcement

This portion of the procedure was started on the day that the last 10 trials with immediate reinforcement in either direction had been given. A switch was set so that the bar withdrew automatically immediately after a correct response, precluding the possibility of the occurrence of additional responses following the one correct response on a given trial. The bar remained in place (inserted in the compartment) during the occurrence of incorrect responses preceding the correct one.[8] Usually a number of incorrect responses were made at first since they were the ones found to have the greatest habit strength in the preliminary training. Simultaneously with the occurrence of the correct response, the Telechron timer started, releasing a food pellet following the proper interval of delay. The animal was allowed 10 sec to consume the pellet and then was returned to its living cage.

To avoid a massing of trials as learning progressed, the inter-trial time for the animals was kept as constant as possible, 10 min elapsing on the average between one trial and the next for a given animal.

If an animal failed to make a response (correct or incorrect) within a five-minute period on three successive trials, it was considered 'extinguished' and training was halted.

Forty training trials were given to each rat for three successive days (excepting those animals that extinguished), making a total of 120 trials involving the differentiation of a response under one condition of delayed reinforcement.

Results

Equivalence of Experimental Groups

The six experimental groups corresponding to the six conditions of reinforcement studied can be compared on the basis of the last two trials of the preliminary bar-training procedure. Table 1 shows the mean response latency for these trials for each group of 25 animals and the standard deviations for the means. Table 2 shows the

TABLE 1

Equivalence of the Six Experimental Groups in Terms of the Latencies for the Last Two Responses of the Initial Training Series (Immediate Reward for Either Direction of Push)

The means and standard deviations are given here. Critical ratios for the differences are given in Table 2. $N = 25$ animals in each group.

	Delay Group					
	0 sec	2 sec	5 sec	10 sec	20 sec	30 sec
Mean latency	3.97	2.97	3.32	3.59	3.80	3.16
σ	4.51	2.92	3.13	4.23	3.93	3.52

[8] In the original plan of the experiment, provision was made for the bar to withdraw following both correct and incorrect responses. With this arrangement the only difference resulting between responses was the delivering of food following the correct response. However, preliminary experimentation indicated that a differentiation of responses under these conditions is very difficult, several animals having a food delay of two or five sec showing no signs of learning the correct response. With some animals several hundred incorrect responses were made in succession without the giving of any food reward. It is believed that very strong position habits and the possibility for spontaneous recovery of extinction effects during the period between trials account for this failure to learn. Experimenters working with discrimination problems (9, 11, 12) have pointed out essentially the same phenomenon.

critical ratios of the differences between mean latency for paired delay groups, indicating that at this stage of training there were practically no differences between groups with regard to the time required for a bar response to occur after the starting gate had been raised.

TABLE 2

CRITICAL RATIOS FOR THE DIFFERENCES BETWEEN PAIRED GROUPS FOR THE LAST TWO PRELIMINARY TRIALS

P equals the probability of obtaining differences of the magnitude shown in Table 1 on the hypothesis that the groups were drawn from the same population.

Paired Delay Groups	C.R.	*P*
0–2	.93	.18
0–5	.60	.27
0–10	.31	.38
0–20	.11	.46
0–30	.71	.24
2–5	.41	.34
2–10	.61	.27
2–20	.62	.27
2–30	.21	.42
5–10	.26	.40
5–20	.36	.36
5–30	.17	.43
10–20	.18	.43
10–30	.39	.35
20–30	.46	.32

Percent of First Responses Correct

This is considered by the writer to be one of the measures of habit strength available from the data which can be used to show a picture of the effect of increasing delay between a response and reinforcement on the differentiation of that response from the other available bar response. It must be remembered that on the basis of the preliminary bar training the least preferred of the two available bar responses was selected as the one to be correct under conditions of delayed reinforcement. Hence, it may be expected that at the beginning of this delay procedure the now incorrect response would show a much greater habit strength than would the correct response. As learning of the correct response progresses, it will be found to occur more and more frequently as the first and only response for a single trial until finally, if the correct response is fully learned, it will always occur and incorrect responses will be eliminated completely.

The results of the learning of this single response under conditions of simultaneous reinforcement (zero delay) are given in Table 3, and the values from this table are presented graphically in Fig. 1.[9] Ex-

[9] The values in Table 3 represent pooled results for successive 10 trials in the experimental boxes for the 25 animals in the zero-delay group. Pooling was resorted to in order to minimize the effects of chance fluctuation and to make it possible to secure a median value for the group

TABLE 3

MEDIAN AND MEAN PERCENTAGES OF SUCCESSIVE 10 TRIALS WHICH WERE WITHOUT ERROR, *i.e.*, INVOLVED NO INCORRECT RESPONSES

The standard deviations of the mean values are also presented. Each value represents 10 trials for each of the 25 animals in a group.*

Practice Trials	Group														
	0-delay			2-sec delay			5-sec delay			10-sec delay			20-sec delay		
	Md.	*M*	σ	Md.	*M*	σ	Md.	*M*	σ	Md.	*M*	σ	Md.	*M*	σ
1–10	30	32.4	21.1	10	15.6	17.9	20	26.0	21.1	10	15.4	24.7	10	17.5	27.1
11–20	80	69.6	28.5	40	43.6	17.9	50	43.2	19.2	20	31.0	20.1	20	25.0	23.5
21–30	90	83.2	25.7	70	68.4	20.8	40	45.2	24.4	40	36.5	20.2	40	31.9	26.2
31–40	100	95.2	20.9	80	78.8	23.0	60	67.2	22.4	50	52.0	25.3	50	45.7	28.2
41–50	100	96.2	21.4	90	88.4	24.0	70	70.4	29.1	60	62.5	31.2	40	44.2	33.5
51–60	100	95.2	23.5	100	93.6	21.9	80	76.4	23.3	60	69.0	27.1	60	53.1	30.1
61–70	100	97.6	21.2	100	95.2	21.3	90	88.0	19.5	80	73.3	24.4	70	58.8	28.6
71–80	100	100.0	20.4	100	96.4	21.1	100	86.8	25.2	90	82.5	25.4	60	65.9	29.0
81–90	100	98.0	20.6	100	97.2	20.8	90	86.0	22.5	90	84.8	25.5	60	66.3	27.2
91–100	100	100.0	20.4	100	99.2	20.4	90	88.0	23.3	80	82.8	22.7	70	73.1	26.1
101–110	100	98.4	20.4	100	98.0	20.6	100	88.8	23.0	90	85.0	21.8	80	75.5	24.2
111–120	100	97.2	20.6	100	100.0	20.4	100	94.4	21.1	90	84.0	22.1	60	72.0	25.7

* Since the score for each animal was the number of trials out of 10 which were without error, the percent of correct trials was easily determined by multiplying the score by 10. The score for the middle animal of a group of 25 animals (arranged in ascending order) multiplied by 10 represents the median percent of trials which were correct. The mean score of the 25 animals multiplied by 10 represents the mean percent of the total 10 trials which were correct. The 30-sec delay group is not included here for reasons which will be discussed later.

amination of Fig. 1 discloses that, when reinforcement occurs simultaneously with the response, learning is quite easy for the rats. Thirty-two percent of the first 10 trials were correct on the average for this group of animals and 95 percent of the trials between the 30th and 40th resulted in the first and only response being correct. It is obvious from Fig. 1 that under these conditions animals can learn the correct response to a nearly perfect degree.

Referring to the mean values presented in Fig. 2 it is apparent that although the 2-sec delay group eventually reaches almost perfect performance the rate with which perfection is approached is somewhat slower than for the animals having no food delay.

With the 5-sec delay group (Fig. 3) there is evidence of a somewhat slower rate of increase for the practice curve and, in addition, there are indications that the correct response never reaches as high

results. These medians were found by calculating for each animal the total number of correct first responses per successive 10 trials and then finding the median for the distribution of all 25 animals. Even with this pooling the crude median can be found only for successive 10 percent intervals up to 100 percent. Because of the grossness of these median values the curve based on them lacks the smoothness found for mean values and consequently comparison of the different experimental groups will make use of the means.

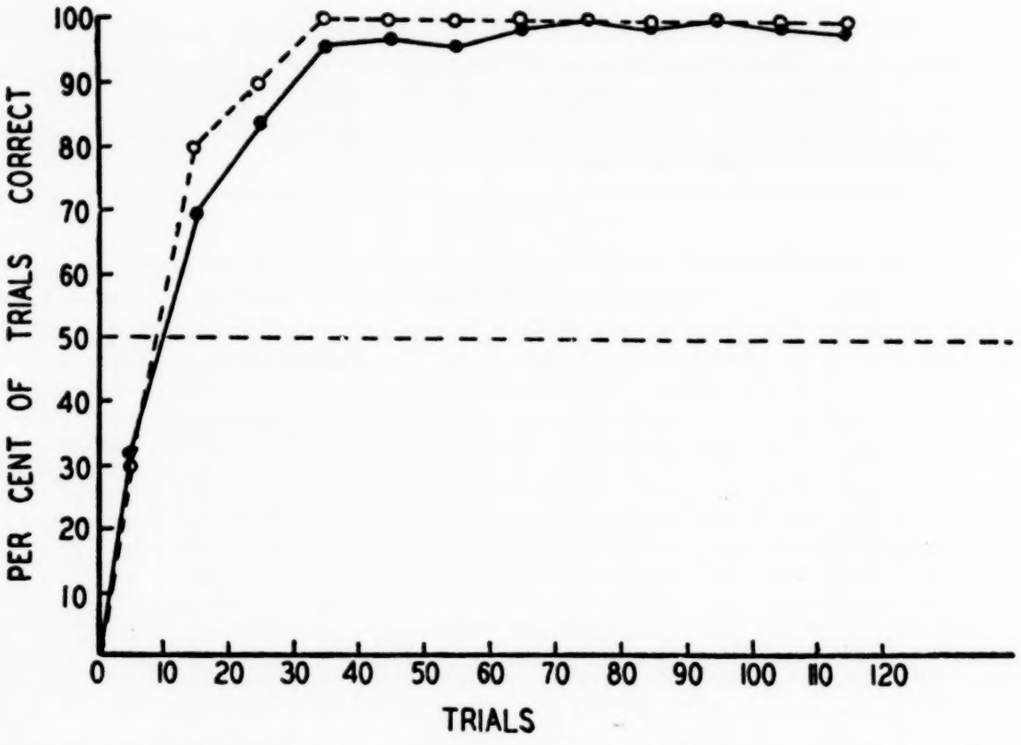


FIG. 1. Plot of the median (dashed line) and mean (solid line) values for the zero-delay group based on the values presented in Table 3. The ordinate represents the percent of trials on which no errors were made. The horizontal line at the 50 percent point represents equal occurrence of correct and incorrect trials.

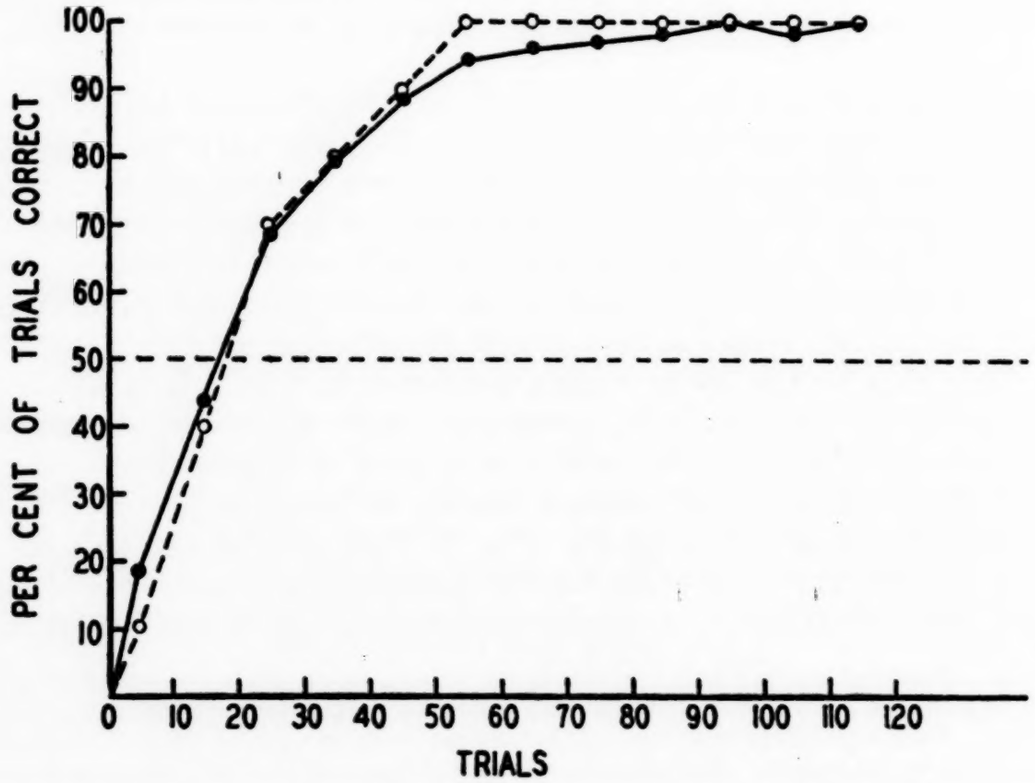


FIG. 2. Plot of the median (dashed line) and mean (solid line) values for the 2-sec delay group based on the values presented in Table 3. (See legend for Fig. 1.)

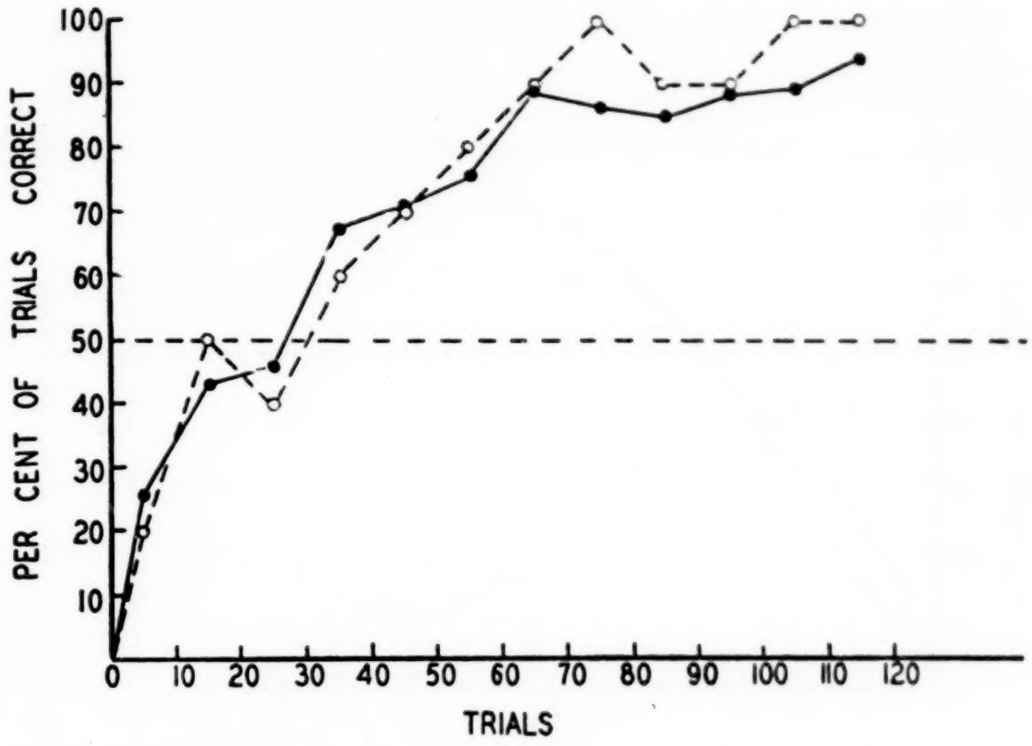


FIG. 3. Plot of the median (dashed line) and mean (solid line) values for the 5-sec delay group based on the values presented in Table 3. (See legend for Fig. 1.)

a percentage of occurrence as those for the zero and the 2-sec delay groups.

The 10-sec delay group (Fig. 4) shows practically the same rate of increase as does the 5-sec group in the early stages of training but the level of the curve at the end of training is lower than that of the latter, indicating, perhaps, that the level of perfection ultimately reached with this (10 sec) amount of food delay is still lower than the level which can be attained with delay of 5-sec.

With the 20-sec group a factor is present which was not found with the groups having shorter delay intervals. With the 0, 2, 5, and 10-sec groups all of the animals completed the 120 training trials, making responses (incorrect and/or correct) with every trial. In the 20-sec group, however, three animals extinguished before the training was completed; that is, three animals ceased to make bar responses. One animal extinguished after five trials, another after 74 trials, and the third after 103 trials. The fact that these animals extinguished makes it more difficult to plot practice curves for the group because it adds a significant departure from the dimension of correct versus incorrect responses which was used with the shorter delay groups. A less desirable way of plotting the practice curve for this group was carried out but it probably yields a fairly typical curve for the group as a whole. The assumption was made that the three animals ex-

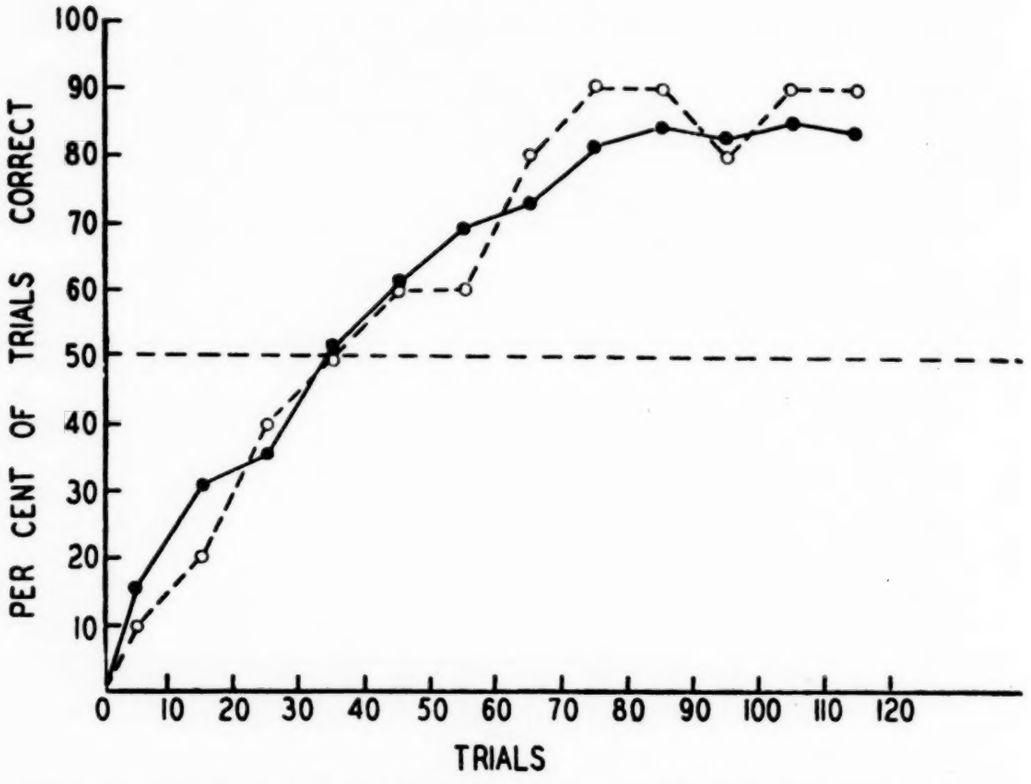


FIG. 4. Plot of the median (dashed line) and mean (solid line) values for the 10-sec delay group based on the values presented in Table 3. (See legend for Fig. 1.)

tinguished because of reduced motivation resulting from the long food delay following correct responses and the extinction effect of incorrect (unreinforced) responses. These animals were discarded from the distribution and to equalize this unbalanced result the three animals which appeared to be the fastest learners (ergo, most intelligent and most highly motivated) were also excluded from the distribution. The mean and median practice curves for the remainder of the group (19 animals) are shown in Fig. 5. Examination of this practice curve discloses that the rate of rise is considerably slower than for the other three groups and also the level of performance reached after 120 practice trials is the lowest of all. It must be recognized that with this particular measure a value of 50 percent represents an equal number of correct and incorrect responses and that the level of performance must be substantially above 50 percent correct in order to show that differentiation of the correct response (*i.e.*, learning) has occurred.

With the 30-sec delay animals the factor of extinction was so great that it precluded the possibility of plotting practice curves for the group. Nine of the 25 animals ceased to make bar responses before the completion of the 120 training trials. The implications of this increasing tendency to extinguish with longer delays will be discussed later.

Data were also available for the number of errors per trial and the percent of *responses* which were correct. These are probably highly correlated with the measure of learning already presented, and the conclusions to be drawn from them are virtually the same.

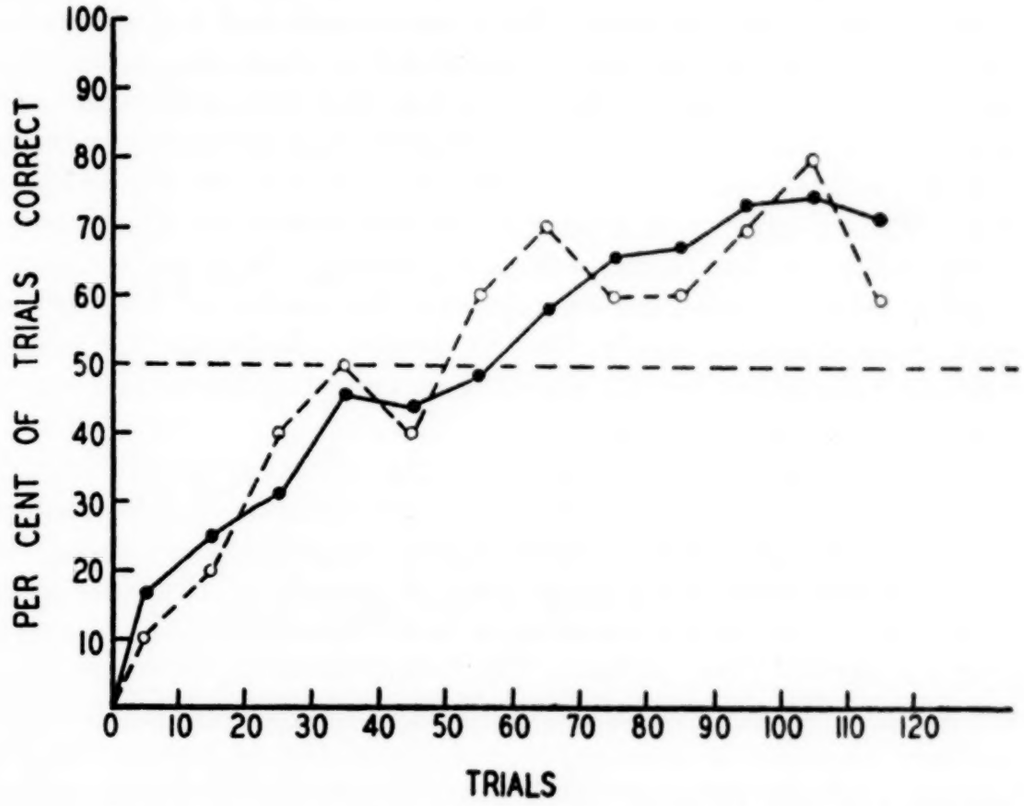


FIG. 5. Plot of the median (dashed line) and mean (solid line) values for the 20-sec delay group based on the values presented in Table 3. (See legend for Fig. 1.)

INTERPRETATIONS AND CONCLUSIONS

It has been reported previously (**8**) that both the rate of acquisition of a habit and the ultimate limit which the strength of the habit can attain are functions of the length of the delay interval. The present results for the percent of trials beginning with a correct response (Figs. 1, 2, 3, 4, and 5) indicate that with this measure, under the conditions of the present experiment, the rate of acquisition is an inverse function of the length of the interval between the correct response and the reinforcement. Further examination of these figures discloses that both the zero-delay and the 2-sec delay groups almost reach perfection as practice trials are increased and that the animals of the 5-sec delay group probably would reach about 95 percent correct responses with continued training. A marked decrease in the ultimate strength of the habit is not noticeable until the delay period is as long as 10 or 20 sec (Figs. 4 and 5). It is believed that the fail-

ure of this measure to discriminate between the shorter delay groups in terms of the ultimate limits attained can be attributed to a combination of two factors. It is evident that animals could not exceed 100 percent perfect trials. Therefore, if the task required of them is sufficiently simple for this level to be reached eventually with a delay as long as two or five seconds, this measure will fail to differentiate the groups. The second factor involved is that the preliminary training presumably gave the required response sufficient habit strength both through direct reinforcement and through generalization from the reinforced incorrect response for it to be learned to an almost perfect degree even when the reinforcement for that response was delayed in the later differentiation training. It is not considered that these results invalidate in any way the results of the previous delayed reinforcement study by the writer. Instead, the present situation is believed to be too complex to yield similar results regarding the limits of habit strength.

Since the task to be learned was the selection of the initially weaker of two available responses, at the expense of the stronger, through differential reinforcement under conditions of delay, the correct response must occur more than 50 percent of the time as the first response of the trial if learning is to be demonstrated. In terms of practice curves then, animals that were learning the correct response would evidence the fact by showing an increasing tendency for the correct response to occur over and above the 50 percent level of perfection, and the slope of these practice curves (*i.e.*, rate of rise) at the 50 percent level for the different delay conditions would constitute an indication of the rate of learning under those different conditions.

Simple parabolic functions were fitted in the vicinity of 50 percent correct trials to the curves shown in Figs. 1–5, and tangents to these curves at the 50 percent level were taken as indications of the rate of change at that point. Table 4 shows the tangent values for the different delay groups, the values being shown graphically in Fig. 6.

TABLE 4

SLOPE AT THE 50 PERCENT POINT FOR THE PRACTICE CURVES BASED ON SUCCESSIVE 10 *Trials*

The slope is given in terms of the tangent to a second degree parabola passing through the 50 percent point.

	Delay Interval				
	0-delay	2-sec	5-sec	10-sec	20-sec
Tangent at 50 percent level	2.92	2.26	1.32	1.16	.59

The equation for a curve fitted to the points in Fig. 6 is:

$$S = 1.6 \cdot 10^{-.15T'} - .043T' + 1.45$$

where S is the slope of the curve and T' is the length of the food-delay period. This is a simple exponential function which approaches a sloping straight line as an asymptote. The curve crosses the axis at 34 sec, which indicates that if the food delay were this long the correct response could not be learned.

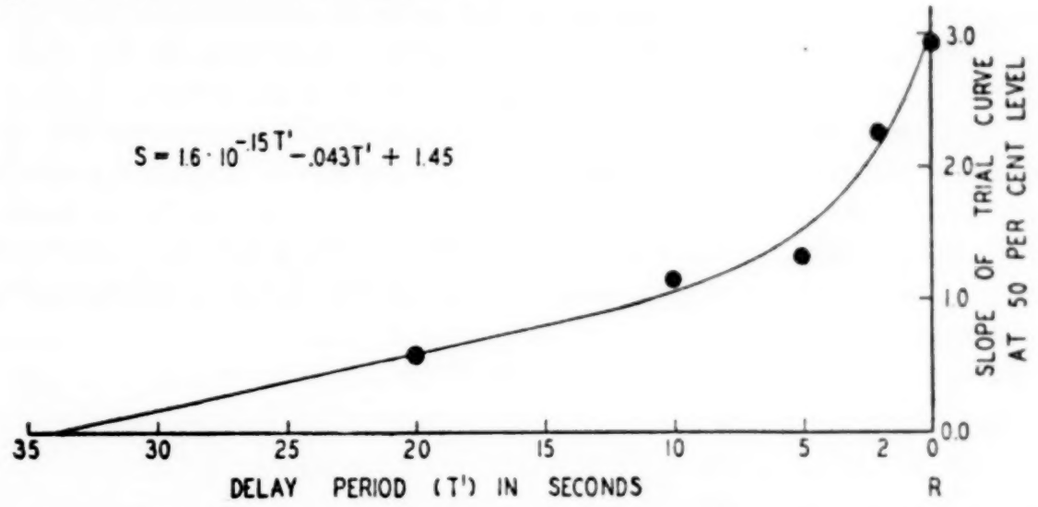


FIG. 6. Slope (tangent) at the point where 50 percent of the trials were correct. The solid circles represent the tangent values and the solid line is a theoretical curve fitted to these values. The equation for the theoretical curve is also given.

The 30-sec delay animals have been omitted from the discussion thus far because of the tendency in this group for the animals to extinguish before completion of the 120 training trials. From examination of the 20-sec group (in which three animals extinguished) and the 30-sec group (in which nine animals extinguished) it would seem that the longer the delay interval, the greater the tendency for the animals to cease responding. This increasing latency is not surprising, however, in the light of the probable interaction between the correct and incorrect responses. Since the incorrect responses were not reinforced there was ample opportunity for them to undergo extinction and presumably this extinction effect could generalize to the opposite (correct) response. Now, if the food were given simultaneously with, or shortly after, the correct response the increments of reinforcement to that response might be sufficient to override both the generalized extinction tendency and the inhibition resulting from the occurrence of the correct response itself.[10] It is conceivable, however, that with the longer delay intervals this increment of reinforcement from each trial would be too small to overcome the accumulated

[10] Assuming that the factor of 'inhibition of reinforcement,' described previously by Hovland (3), is operating in the present experimental situation.

inhibition for both responses and thus the animals would cease to make either correct or incorrect responses.

The tendency for the longer-delay animals to cease responding is additional evidence for the conclusion that, with the present conditions and the obtained measures, learning could not take place if the food reinforcement were delayed thirty seconds or longer. This estimated limit of delay is considerably longer, however, than that found in the writer's previous study (**8**). Again it must be emphasized that the correct response in the present experiment was given some training with immediate reinforcement previous to the task of differentiating it from the preferred (incorrect) response. This and the fact that the correct response presumably was susceptible to a degree of positive transfer effect from the incorrect response probably account for the tendency of the differentiation response to be learned under longer delay conditions than found in the previous experiment where learning the bar response was under the delay conditions from the outset.

Theoretical Implications

Two factors in the present experiment demonstrate the extreme importance of secondary or derived reinforcement in the ordinary learning situation. First, the very short limit of delay which will permit learning to take place in the absence of secondary reinforcement tends to show indirectly that this type of reinforcement must be present if the primary reinforcement is very remote from the response which is to be learned. Second, the difficulty in eliminating secondary reinforcement characteristics, with the resulting highly artificial experimental conditions when they are practically eliminated, indicates their generality in the usual learning situation.

It should be emphasized that subgoals, symbolic rewards, and transfer effects from similar responses previously learned all contribute to learning and must be included, in addition to a primary gradient of reinforcement, in any systematic approach to the learning process.

Summary

1. Six groups of albino rats, 25 in each group, were trained in a bar-pressing situation. The bar was capable of being moved in a right or left direction and the task demanded of the animals was the selection of one of these directions of push as the correct response. Provision was made for the bar to withdraw automatically following a correct response.

2. The experimental variable differentiating the six groups was the interval of time between the correct response and food reinforcement for the response. Delay intervals used, corresponding to the six groups, were zero, 2 sec, 5 sec, 10 sec, 20 sec, and 30 sec.

3. All of the animals were given 60 training trials on the bar responses with reward being given immediately for either direction of push.

4. Differential reinforcement was then given to the animals with the preferred response (on the basis of 3 above) being incorrect and with the alternative response being reinforced following the delay interval for the experimental group. A correction technique was employed by which each trial ended with a correct response.

5. Practice curves for five of the six experimental groups show the percent of trials made without error. The limits of habit strength in terms of perfection of the response differentiation with practice does not discriminate between the zero and the 2-sec delay groups. Additional delays (5, 10 and 20 sec) result in a decrease in the level of perfection.

6. The rates of rise at the level of 50 percent perfection are given for the different delay groups and demonstrate that the rate of habit acquisition is an inverse function of the delay-of-reinforcement interval, *i.e.*, the greater the delay the less the rate at which the habit is acquired.

7. The results indicate that if the food were delayed longer than 30 sec the correct response would not be learned.

8. Some theoretical implications, particularly with regard to the importance of secondary reinforcement, are given.

(Manuscript received July 1, 1942)

References

1. Bugelski, R. Extinction with and without subgoal reinforcement. *J. comp. Psychol.*, 1939, **26**, 121–134.
2. Ellson, D. G. Quantitative studies of the interaction of simple habits. I. Recovery from specific and generalized effects of extinction. *J. exp. Psychol.*, 1938, **23**, 339–358.
3. Hovland, C. I. 'Inhibition of reinforcement' and phenomena of experimental extinction. *Proc. nat. Acad. Sci.*, 1936, **22**, 430–433.
4. Miller, N. E. & Dollard, J. *Social learning and imitation.* New Haven: Yale University Press, 1941.
5. Mowrer, O. H. Motivation and learning in relation to the national emergency. *Psychol. Bull.*, 1941, **38**, 421–431.
6. Perin, C. T. Behavior potentiality as a joint function of the amount of training and the degree of hunger at the time of extinction. *J. exp. Psychol.*, 1942, **30**, 93–113.
7. Perin, C. T. A quantitative investigation of the delay-of-reinforcement gradient. *J. exp. Psychol.*, 1943, **32**, 37–51.
8. Perin, C. T. A quantitative investigation of the delay-of-reinforcement gradient. Ph.D. dissertation, Yale University, 1942.
9. Riesen, A. H. Delayed reward in discrimination learning by chimpanzees. *Comp. Psychol. Monogr.*, 1940, **15**, No. 5, pp. 54.
10. Williams, S. B. Resistance to extinction as a function of the number of reinforcements. *J. exp. Psychol.*, 1939, **23**, 506–522.
11. Wolfe, J. B. The effect of delayed reward upon learning in the white rat. *J. comp. Psychol.*, 1934, **17**, 1–21.
12. Wood, A. B. A comparison of delayed reward and delayed punishment in the formation of a brightness discrimination habit in the chick. *Arch. Psychol.*, 1933, **24**, No. 157, pp. 40.

SOME PARALLELS BETWEEN PUPILLARY 'REFLEXES' AND BRIGHTNESS DISCRIMINATION

BY S. HOWARD BARTLEY [1,2]

Washington University School of Medicine

Among the major problems which confront those who deal with vision is the question whether recordable physiological parallels to *brightness discrimination* can be found in the visual pathway. Up to now no significant analogies have been recognized, for on almost every hand measurable features of physiological response have seemed to vary only with the total flux of light employed in stimulation, whereas brightness discrimination, a sensory response, has varied all the way from dependence upon total flux to dependence upon flux per unit area, and finally almost to independence of the amount of light used.

In pupil reactions to light we have a physiological response which may possibly display some of the properties which are sought for. This was suggested by the fact that in the production of pupillary constriction area and intensity are far from reciprocal. Be the significance of this as it may, investigation of the pupillary reaction to light was undertaken to see how closely it might compare quantitatively to brightness discrimination in a number of employable cases.

Brightness discrimination is measured only by the comparison of one stimulus area with another, as for example, when the intensity of one of two adjacent areas is adjusted so that the two appear equally bright. Brightness discrimination is a matching process, for when the areas differ in physical intensity the observer cannot dependably judge the amount by which one area is brighter than the other.

The pupillomotor response to a single stimulus area can be measured without regard to the brightness of any other area, or set of conditions. In this sense it is absolute and not relative in its measurement.

Pupillary comparisons may be made, however, by measuring the pupils of the two eyes simultaneously and independently stimulated. When the two eyes are confronted with equal test objects, their pupils ought to constrict equally. The well known 'consensual light reflex' demonstrates that light falling on one eye induces, at least

[1] This work was done under a Grant-in-Aid from the Rockefeller Foundation and from the Committee on Scientific Aids to Learning, National Research Council.

[2] I wish to thank Dr. George H. Bishop for the valuable suggestions he offered during the course of the work.

under some conditions, constriction in the other, so that light falling on the two eyes, whether equal in amounts or not, may possibly be reflected in the constriction of each of them. A sort of equilibrium analogous to brightness discrimination may take place.

For purposes of testing the possible outcome, it was first assumed that when the two pupils constrict equally, the stimulus-objects confronting the two eyes are effectively equal just as when the two objects *look* equal in brightness, regardless of whether they differ in other ways or not. But to anticipate in preparing the reader for the actual findings, the nature of some of them must be revealed here. First of all, the above stated simple notion of the stimulus conditions producing pupil equality was not sufficiently inclusive of the variety of conditions which were found to produce it. In the second place, pupil responses under some conditions paralleled sensory response to an extent enabling reactions to differentiated fields confronting even a single eye to be measured.

Apparatus and Method

The essentially different stimulus arrangements employed should be recognized before we proceed any further. They are as follows: (1) Each eye independently confronted with one of two equal-sized discs, and the intensity of one varied from dim to bright while the other is kept at the maximum brightness of the variable; (2) The use of two *unequal*-sized discs, the intensity factor handled precisely as in case one. In this general set of conditions there are three possibilities; namely, three sets of relationships between the small and large discs. The first (2a) is disc 1 (small area) and disc 2 (medium), the two having an area ratio of 1 to 4. The second (2b) is disc 2 and disc 3 (large area), with a ratio of 4 to 28 (1 to 7). The third (2c) is disc 1 and disc 3 with a ratio of 1 to 28. In each case the larger disc was varied in intensity. In essence 2a represented two discs not greatly different in area; 2b, two discs differing much more greatly in area. The last case (2c) involves not only *difference* but also the possibility that the greater of the two discs by its preponderance will impose its effects on the smaller so as to make the pupils act more nearly alike. (3) The illumination of only one eye. That is to say, a still wider discrepancy between stimulation of the two eyes—to the point where the one eye receives none. (4) The stimulation of one eye with a disc whose intensity is varied from dim to bright, following which the corresponding area in the other eye is also stimulated by a similar disc whose intensity likewise begins at a very low level and is increased to the level of the first. (5) The use of the same procedure on non-corresponding points of the two eyes. (6) The use of the same procedure on two areas of *one* retina.

The apparatus used in the experiments of this study was the same, except for slight modifications, as that described in a paper soon to be published (Bartley). Hence it need only be stated here that pupil size was determined by motion picture photography using infra-red radiation. In the first three sets of experiments, readings were taken both with ascending and descending shifts in intensities of the variable, the two series being averaged to give the final measure. Various amounts of preadaptation were tried to find one giving constancy of result, hence various durations, from the time required to turn off the room light and start the camera motor to 15 min. in darkness, were tried. The interval between readings was also varied from the time needed to manipulate the two switches mentioned, to 5 min. of darkness.

Results

Experiment 1.—The outcome of using two equal-sized discs, one of which is varied in brightness, is shown in Fig. 1. Although the

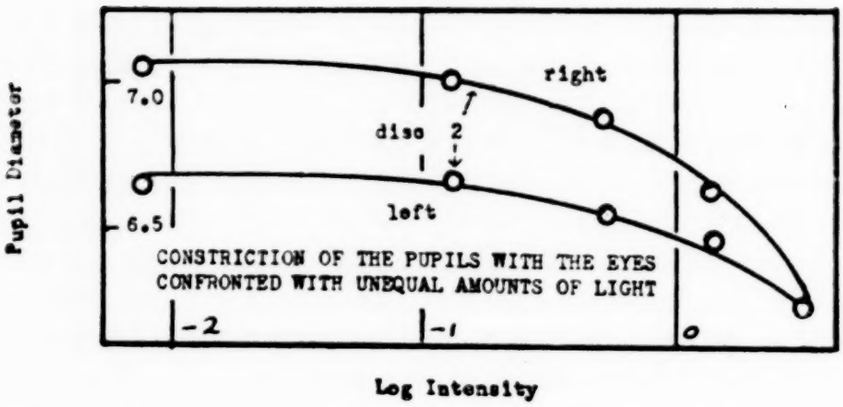


FIG. 1. Differential constriction of the two pupils when two equal-sized discs are independently presented to the eyes. At first only one is bright, the other being progressively brightened from an initial low level. (Abscissæ in all figures, c./ft², ordinates, millimeters.)

readings were made both by decreasing the variable from a brightness matching the fixed level of the other disc to the final minimum, and increasing it from this minimal level to final equality, the two sets of readings were averaged and plotted in the increasing order. From the figure it will be noted that as the dimmer of the two discs was increased in intensity, both pupils constrict, rather than only the pupil of the eye variably stimulated. This is a demonstration of the consensual light reflex as it occurs under this set of conditions. As expected, the two pupils reach equality in diameter when the two discs become equally bright. Up to this point, the pupil of the variably stimulated eye seems to undergo about twice the reduction in diameter the opposite eye does.

Experiment 2a.—When the two discs are unequal in size, and the intensity of the larger reduced to minimum, the two pupils were found to be unequal in diameter (Fig. 2). The larger disc produced the smaller pupil, although the difference between the two was not great.

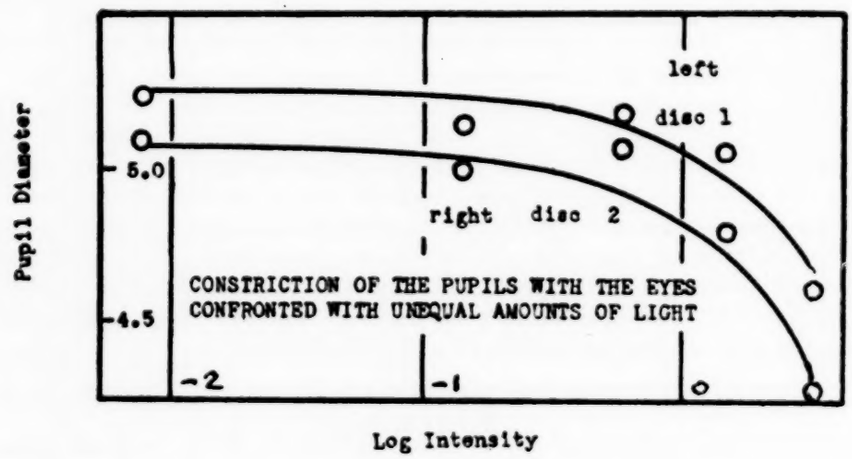


FIG. 2. Pupil response to unequal-sized discs, the experimental procedure being the same as in the last case. Note that the two curves diverge as the dimmer disc is raised in intensity.

As intensity is increased, the pupil of the eye directly concerned constricts more than the other one, so that the curves diverge. This is as might be expected, for a large bright disc differs more greatly from the small bright one in total flux than when the large disc is dim.

Experiment 2b.—The combined area of the two unequal-sized discs in this experiment, and the *difference* in their areas, were greater than in the last. In this case the pupil of the eye stimulated by the larger and dimmer disc was constricted less than the other one at the bottom of the intensity scale, but as intensity was increased constriction was so great that the two curves crossed (Fig. 3). The fact that the curves

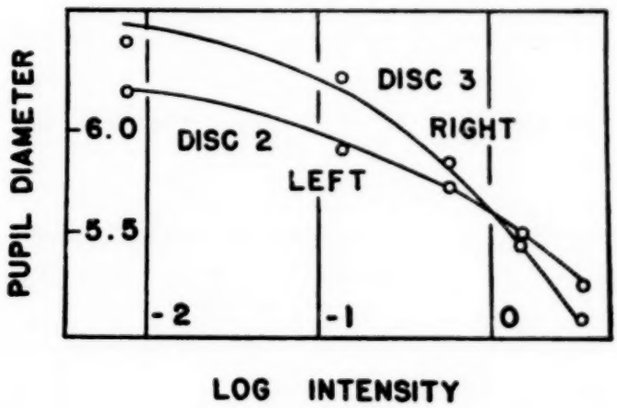


FIG. 3. Pupil response in another situation when the two discs are unequal. Here the curves cross, owing to the fact that the large dim disc, which was initially less effective than the other, finally becomes the more effective when intensity is increased.

cross indicates that the relative effectiveness of the two discs reverses, from the larger being the less effective to being the more effective. The point of crossing shows the region of equality, and that, although area and intensity compensate for each other, the two are not simple reciprocals.

Experiment 2c.—In this experiment the two discs were also unequal in size, the modification from the last experiment being that the relative areas are still more diverse. In this case the curves representing the two pupil diameters lie fairly close together, even at their greatest divergence (Fig. 4). At the lowest intensity for the larger but dimmer disc, they are identical, but as intensity is increased to equality with the other, some divergence takes place. Since the areas of the two stimulus discs were relatively more different in this case than in the previous two, we might have expected that the pupil sizes resulting would have been *more diverse*, rather than less. That the two curves for the two eyes slope at different rates makes the results here like those in *2a* and *2b*, but the fact that the pupil diameters tend to be more nearly equal than in the two previous cases demonstrates the entrance of another influence. This is that when one

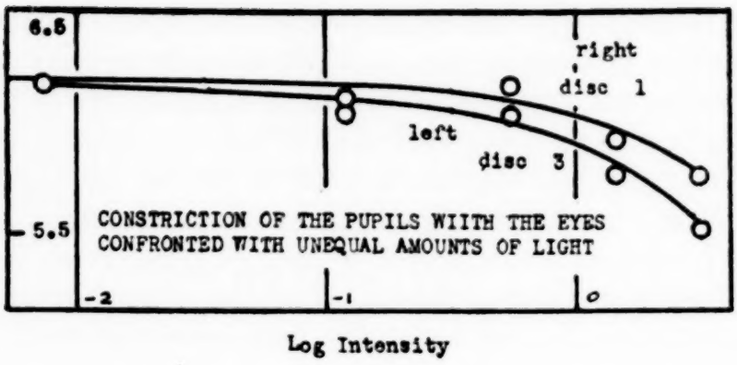


FIG. 4. Outcome in another case of inequality in area of stimulation. (See text.)

stimulus area is very small in comparison to the other, the eye it supplies is dominated by the other, making the two pupils behave nearly alike. The extreme of this principle will be brought out in the next experiment.

Experiment 3.—In this case one eye only received stimulation, for which the variable disc of the previous experiment was used. The results are shown in Fig. 5, wherein it will be seen that the two pupils

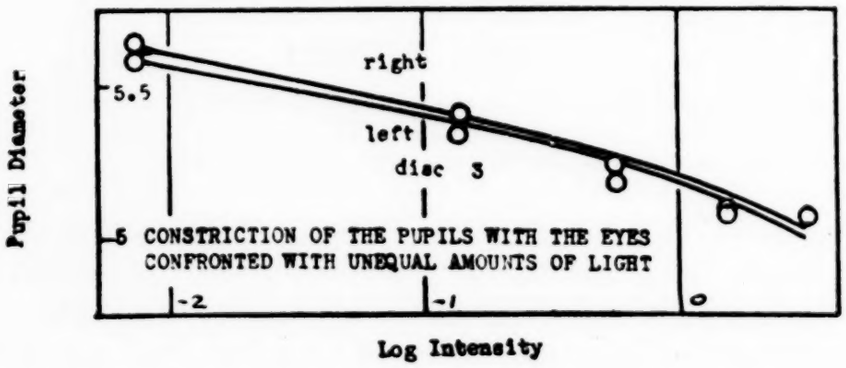


FIG. 5. Pupillary response in both eyes when only one eye is stimulated. This is a demonstration of the consensual light reflex.

behave alike throughout the intensity range, as though they were both receiving equal amounts of light. As was suggested in the previous experiment, when one eye receives a predominant amount of stimulation the other eye is carried along with it. Here we find the consensual light reflex at its maximum effectiveness.

The experiments so far would lead us to believe that not only the absolute area and intensity of each of the stimulus discs but just as greatly their relative values are the basis for the pupillomotor balance reached in the two eyes. In this respect the results are broadly similar to those obtained in brightness discrimination.

Experiment 4.—In this and the succeeding experiments, the disc confronting the one eye (left) was presented first, and its intensity

increased step by step up to maximum where it was held throughout the presentation of the other disc in the same way.

In this experiment, the two discs, presented individually to the two eyes, were placed so close together that relaxation of convergence brought the images of the two discs onto corresponding areas of the two retinæ, and the perception fused into one. That is to say, the observer now had only one disc to deal with subjectively. In Fig. 6, the curve represented by the open circles shows the results.

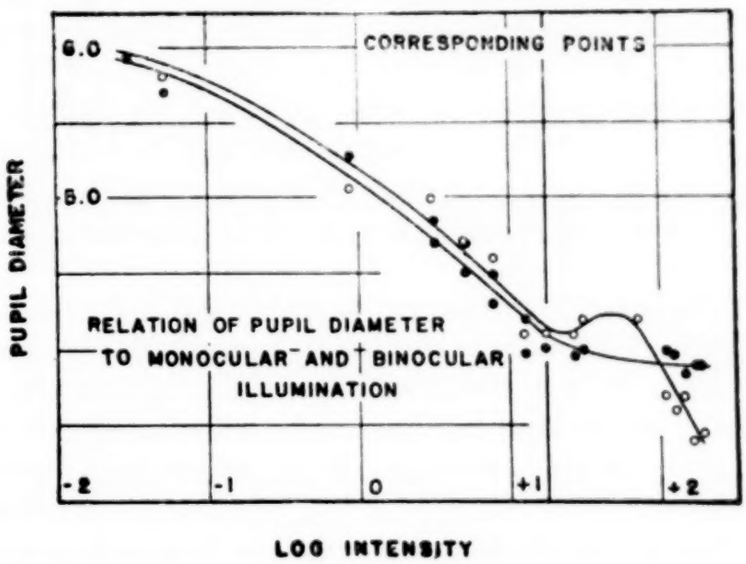


FIG. 6. Pupillary effects in two cases; the one (black circles) in which intensity was progressively increased on a single area of one eye, by first using one disc and then adding light by superimposing the second disc whose intensity was also raised by steps; the other case (open circles) in which the two discs were employed to stimulate the corresponding regions of the two retinæ, the first disc on the left eye, the second on the right.

It will be seen that as the intensity of the first disc was varied, both pupils constrict alike. Similar behavior in the two eyes was true also when the second disc was added, but the pertinent phenomenon to note, however, is that adding the second disc *dilated* the pupils, when dim, but as its intensity was raised, constriction set in again. The curves in the graph again begin to descend.

That the combined effects of the two discs used on corresponding points of the two retinæ are greater than the effect of using the same amount of light in a single area of one retina is shown by a comparison of the present curve with the other one in Fig. 6. This is another example of binocular interaction.

Experiment 5.—The same procedure as in the last experiment was used, except that the discs were too widely separated for adjustment to bring their images onto corresponding points of the two retinæ. Hence they were seen as two discs instead of one. Fig. 7 indicates the outcome. It will be seen that so long as only one disc is used,

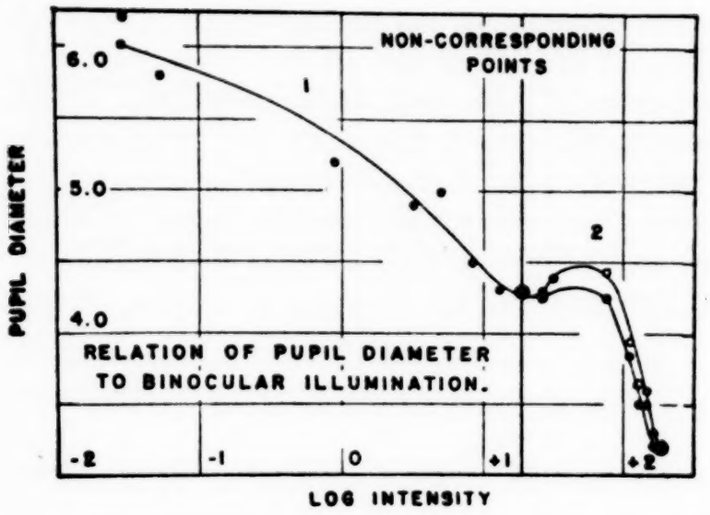


FIG. 7. Pupillary outcome when two non-corresponding areas of the two eyes were stimulated by the same procedure as in the previous experiment.

the pupils behave alike. Both constrict as the single disc is made progressively brighter. As soon as the second disc is introduced the two pupils begin to act differentially. They both begin to *dilate*, the pupil of the newly and less intensely stimulated eye showing the greater amount. Both pupils soon begin to *constrict* again, however, as the intensity of the second disc is made greater and greater. Here, too, the combined stimulation of two eyes is more effective than when the same total amount of light falls on a single area of one eye.

Experiment 6.—In this case the procedure was modified only to provide for the images of the two discs to fall on nearby areas of *one* eye. Again *constriction* is reversed into *dilatation* when the second disc is introduced Fig. 8. The effect seems to be more pronounced here than in the two preceding cases, although it must be said that the experiments were not repeated enough times to reveal precisely

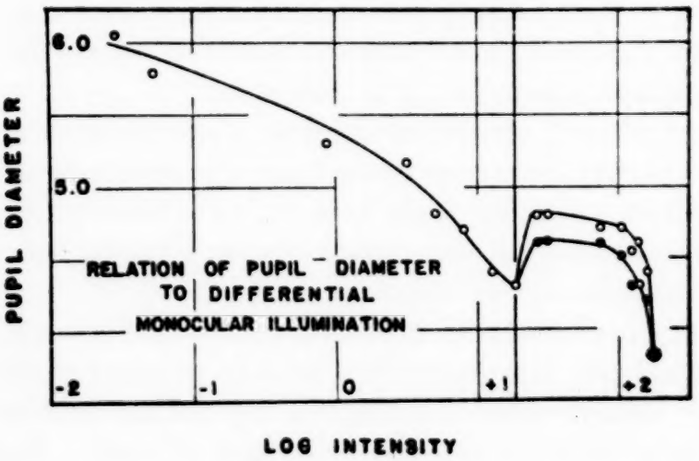


FIG. 8. Results when two nearby areas of the same retina were stimulated by the two discs.

the absolute position of the three sets of curves. Again, also, the combined effects of stimulating two areas with the quantity of light we used was more effective than using it on a single area.

Qualitative Experiments in Brightness Discrimination.—In order to see to what extent the subjective brightness discrimination in each case represented by the preceding three experiments would parallel the pupillomotor results, the stimulus conditions were repeated for the specific purpose of perceptual evaluation. The results of so doing are as follows:

1. When a bright disc stimulates one eye and light from a similar disc is added to the corresponding region of the second retina, the effect is that of reducing the seen brightness of the disc when the physical intensity of the added disc is small in comparison to that of the first disc. But when the intensity of the second disc is made greater and greater, there comes a time when the combined effects produce additional brightness. This can easily be seen by alternately turning on and off the light-source producing the second or superimposed disc, both when it is dim and then when it is brighter and brighter. Alternate adding and removing of the light of the second disc is the best method to make the change in brightness produced apparent when it exists. The results in this case demonstrate Fechner's Paradox.

2. When the second disc is so situated as to be separately seen, its introduction into the visual field produces a reduction in the brightness of the first disc. This is true when the second disc is dim, but when the two discs approach equality in brightness they both look brighter than a single disc by itself.

3. Likewise when a second disc is introduced alongside of a first one, in the field of a single eye, the two changes mentioned in the previous situations will occur. The brightness of the first disc is reduced when the added disc is dim, but is later enhanced when the physical intensities of the two become more nearly equal.

Hence, from these results it is obvious that the sensory and pupillomotor effects parallel each other in quantitative pattern.

Discussion

The findings here bear not only upon problems of *brightness discrimination*, including binocular interaction, but also upon physiological problems of reflexology and peripheral and central nervous organization, including the question of pupillomotor areas of the retina. In fact, not least of all, they bear upon problems of *visual fatigue*. All of these problems are so intimately intertwined that separate discussion of each of them is a bit difficult.

Binocular Interaction.—Binocular vision has given us a variety of phenomena which lend complexity to the subject of brightness discrimination. The phenomena of binocular interaction are, however, but one large part of the total range of brightness discriminatory reactions. The addition of the second eye to the functioning of the first, although it brings in many new phenomena, is not the essential condition for some of the phenomena that might easily be ascribed to binocularity. In some ways the second retina acts as merely an extension of the first. The essentially significant factor in the production of the great diversity in the results obtained in brightness discrimination is the degree and manner of differentiation of the visual field, so that various portions of retinal receiving surface are differentially stimulated, whether they lie in one eye or two.

In due course in the history of the general problem of brightness perception, the question of binocular summation came to the fore. It was well seen as having certain possibilities of analogy to the bilateral reflexes in its basic physiology. The answer, nevertheless, by the then existing authority turned out to be that binocular summation or interaction did not exist, except in the process of making psychologically *one* what was physiologically elaborated as two by the two eyes independently. Prior to this, Fechner's Paradox had arisen and received some attention. It illustrated binocular brightness discrimination in a specific situation, and made plain that binocular 'subtraction' could take place, even though other forms of interaction called summation had not yet been universally found. But the attempt, on the part of some investigators, to deduce binocular interaction from Fechner's Paradox was met by the argument of *psychological* action superimposed upon physiologically independent elaboration by the two eye mechanisms. This concept, based on some experiments not fitted to crucially test the problem of physiological interaction, suspended further progress for a time.

Supplementing the experiments of older investigators who had found some evidence for binocular interaction, and quantifying its existence in experimental situations, De Silva and Bartley (**3**) demonstrated the untenability of the then conventional idea that any binocular interaction was 'psychologically' rather than physiologically brought about. Later Fry and Bartley (**6**) extended the knowledge of binocular vision by suggesting the basis for Fechner's Paradox. The role of spatial differentiation in stimulation began to be demonstrated by showing how borders in a field react upon each other and depress or enhance the appreciation of surfaces.[3]

[3] Lest it be erroneously thought that these authors assigned anything mystical to borders in disregard to organismic conceptions, it must be made clear that the existence of the results they found when borders were placed in certain relations to each other were not assigned to

The present study on the pupil adds another chapter to the understanding of binocular vision, not only by indicating that pupil action 'copies' sensation, but by showing incidentally that light added to the second of two areas perceptually separated in space acts much the same whether the binocular or monocular apparatus is employed in its appreciation. That is to say, in the case in which the second disc is simply superimposed upon the first, by use of the second eye, even though *no new borders* are introduced *in perception*, the results are as if they had been. This is a case in which, though perceptually no *new* borders are set up, certain new border processes are initiated in the nervous system, namely, in the second retina, and in the brain. Those in the retinæ are only able to interact when their effects are transmitted to the brain. Although we cannot outline what must happen, here we have a specific problem, the solution of which will give us a much better understanding of the visual mechanism.

The assurance we gain, however, is that binocular interaction has a physiological basis, and is not, as was insisted upon by Sherrington, dependent upon psychological processes which transcend physiology.

Pupillomotor Areas of the Retina.—The pupillary results show clearly that the pupil action follows much the same pattern as sensation, up to the point where the amount of constriction or dilatation can be determined by the differentiation in the spatial pattern of retinal stimulation. We do not yet know how fine the differentiation must be before the pupil fails to follow it, but it is obvious that the present findings bear upon the long-discussed question of what part of the retina is pupillomotor. From our knowledge now, we would surely have to say that the whole retina is pupillomotor. In saying this, we imply that the center is much more sensitive than the periphery, just as we do with reference to discriminatory sensation. This follows from the fact that we find pupillary action paralleling *form* vision, which itself is far better toward the fovea than elsewhere, the admission of which is not a declaration that the normal individual is blind at the retinal periphery. In describing sensation, we simply state the kind of vision obtainable through stimulation of various parts of the retina, and what this stimulation must be. It would seem that the same essential principles would hold true for handling pupil action which, though slower, is analogous to sensation. That being the case, no longer can it be said that only a small central region

borders as such, but to the processes of which the perception of borders was only the outward symptom. Borders were only usable in speech as the shorthand term to signify the complete neural pattern underlying the results. By the existence of borders in certain spatial relations, it could be predicted that certain areas in certain directions from them would be brighter or dimmer than expected from the norm gained from a geometrically simpler stimulus field. Since we see that pupil response follows, in pattern, the sensory results, we must include the idea that the pupil mechanism is subject to what has been found true about borders.

is pupillomotor; rather, for gross undifferentiated action light falling anywhere on the retina is effective. The concept that springs from our results would seem to bring into line the several divergent types of findings which have appeared in the literature. The principle involved is simply that if gross reactions are called for by undifferentiated stimulation, one result will occur, but if precise reactions from differentiated spatial patterns are taken as the reaction to be measured, another result will occur. All too often these are seen as conflicting, whereas they need not be.

Reflexes.—What has usually been thought of as a reflex (pupillary reflex) is here shown to behave in certain respects both qualitatively and quantitatively like visual perception. Three possible interpretations might be placed on this. It might be inferred: (*a*) simply, that the pupillary reflex is more complicated than has been heretofore apparent; (*b*) that this reflex is under elaborate, if involuntary, control of the higher centers as other reflexes are inferred to be, and that its behavior under given conditions of stimulation reflects the behavior of the more complex centers imposed on it by such control; or (*c*) that so-called reflexes, on the one hand, and the physiological counterparts of sensory discrimination on the other, are not essentially different in kind, being both the complex functioning of networks of neurons. Consciousness arising from the one set of neurons and not from the other might be quite incidental to the question of their pattern of behavior.

The latter possibility has been suggested above, if only by inference rather than demonstration; that is, it is possible that sensory discrimination via a cortical mechanism, and pupillary 'discrimination' by a reflex mechanism in the 'lower' centers, act in many respects similarly, differing chiefly in speed of response. For purposes of analysis, the question might be put differently; taken by itself, and disregarding phenomena of sensory discrimination, is there anything about the behavior of the pupillary response here demonstrated that could not be interpreted in terms of the properties of reflexes in general: facilitation, summation, reciprocal inhibition, etc.? In terms of integrated nervous action rather than in terms of sensory discrimination, or the supposedly unique kind of neural organization underlying it, can corresponding phenomena be demonstrated on a known reflex level? This might be tested either by investigating the pupillary 'reflex' in a decorticate animal, or by testing the behavior of other known muscle reflexes activated by patterns of stimulation analogous to the two-area stimulation employed here. For instance, limb reflexes subject to multiple avenues of activation from either side of the body might show complication of interaction between two stimuli comparable to the 'discriminating' action of the pupillary response.

Visual Fatigue.—Eye muscles when operating in the usual way are capable of prolonged comfortable actions. But we know that more or less discomfort and inefficiency are common in a great many cases. In discomfort and fatigue there are two components, the attitudinal and the specific muscular component in the eyes themselves. In the present case we are interested in the latter, and have found that the pupillary and the extrinsic eye muscles contribute to the uncomfortable feelings when used in certain ways. A paper about to be published demonstrates certain activity of this kind. The present paper, in showing the intimate connection between pupillary muscle activity and perception through the demonstrated parallelism in action, indicates that the pupillary mechanism is subject to more minute variations of influence from the afferent visual arc than previously supposed. These variations in being the same kind as reflected in perception suggest that the pupillary mechanisms can become involved in the same conflicts as perception. Conflicts in muscle activity induce feelings of tension and strain which become components of visual fatigue.

Summary

The pupils of the two eyes were measured by infra-red photography during response to independent stimulation bydiscs of various sizes and intensities. In certain cases, ascending steps in total quantity of light were presented to the visual mechanism, the first part of the series provided by one disc, and the last part by the other.

In the simpler situations, the two eyes responded more or less on the basis of total flux, intensity and area not being equally effective. When the quantity of stimulation on the one eye was very small, or stimulation was entirely absent, the pupil followed the other pupil precisely. This was a demonstration of the maximal consensual reflex.

In the complex cases (mentioned in the last part of the first paragraph, above) the behavior of the pupils paralleled the outcome characteristic of brightness discrimination under such conditions. Their quantitative pattern of action was parallel to that of brightness discrimination in demonstrating Fechner's Paradox.

The relation of the findings to binocular interaction, reflex action, pupillomotor areas of the retina, and visual fatigue was discussed.

(Manuscript received August 26, 1942)

References

1. Bartley, S. H. *Vision: a study of its basis.* N. Y.: Van Nostrand, 1941, pp. xv + 350.
2. Bishop, G. H., & O'Leary, J. Potential records from the optic cortex of the cat. *J. Neurophysiol.*, 1938, 1, 391–404.

3. De Silva, H. R., & Bartley, S. H. Summation and subtraction of brightness in binocular perception. *Brit. J. Psychol.*, 1930, **20**, 241–250.
4. Ferree, C. E., Rand, G., & Harris, E. T. Intensity of light and area of illuminated field as interacting factors in size of pupil. *J. exp. Psychol.*, 1933, **16**, 408–422.
5. Fry, G. A. Relation of border contrast to distinctness of vision. *Psychol. Rev.*, 1931, **38**, 542–549.
6. Fry, G. A., & Bartley, S. H. The brilliance of an object seen binocularly. *Amer. J. Ophthalmol.*, 1933, **16**, 687.
7. Graham, C. H. An investigation of binocular summation. *J. gen. Psychol.*, 1930, **3**, 492–510.
8. Helmholtz, H. v. *Handbuch d. Physiol. Optik*, (3rd ed.) II, 278.
9. Hess, C. The extent of the pupillo-motor areas of the retina and the pupillo-motor receiving organs. *Arch. Ophthalmol.*, 1909, **38**, 131–146.
10. Luckiesh, M., & Moss, F. K. Area and brightness of stimulus related to the pupillary light reflex. *J. opt. Soc. Amer.*, 1934, **24**, 130–134.
11. Sherrington, C. S. *Integrative action of the nervous system.* London, 1906, 371–375.

A STUDY OF FRUSTRATION IN NEWBORN INFANTS*

BY DOROTHY POSTLE MARQUIS

Yale University

PROBLEM

The present study was designed to investigate the reactions of newborn infants to frustration in the feeding situation. It is part of a general program of investigation to study the adaptation of the newborn infant to the cultural demands of his post-uterine environment. The author in another study (**6**) has shown that rigid adherence to the four-hour feeding schedule frequently imposed on the newborn infant is in itself a frustrating situation, since successful adjustment to such a schedule is a learned response involving, in the beginning, at least, interference with the established food-seeking goal responses present at birth.

Sears and Sears (**7**) in one infant of 23 weeks studied frustration under different strengths of the hunger drive by withdrawing the bottle after different amounts of milk had been consumed, and found that the latency of the crying response following this withdrawal varied directly with the amount of milk which had been consumed. They concluded (**7**, p. 00) that in this infant "the strength of a frustration-reaction varies directly with the strength of instigation to the frustrated goal-response." The term 'frustration' in the present study is used in the sense described by Dollard, Doob, Miller, Mowrer, and Sears (**2**, p. 11), as "that condition which exists when a goal response suffers interference." The specific problems which the study was designed to investigate were the nature of the reaction-patterns present in the neonate to a type of frustration common to most bottle-fed infants, the variation of reaction-patterns from infant to infant, and the factors determining individual variations if they are found to occur.

SUBJECTS AND METHOD

Subjects were 7 newborn full-term infants of the Department of Pediactrics of New Haven Hospital. They were unselected except for the factor of bottle-feeding and normal physiological condition upon examination after birth. Birthweights varied from 2250 to 4310 grams. All infants were on a four-hour feeding schedule, the amount of formula at each feeding being determined by the hospital standards for the correct number of calories per kilo of body weight, and to some extent by the infant's appetite. Attempt was made to observe each infant for two

* From the Yale Institute of Human Relations and the Department of Pediatrics, Yale School of Medicine.

feedings each day during the ten-day period of hospitalization. Because, however, of occasional elevated temperatures or other unpredictable events, daily observations were not complete for all infants.

The experimental procedure was as follows. The infants were brought from the nursery to an experimental room for feeding and placed in a bassinet equipped for automatic registration of general bodily activity (6). After a five-minute interval the experimental record was begun. The experimental period consisted of 6 periods of recorded observations distributed as follows: (1) A pre-feeding period of 90 sec., after which the infant was permitted to take one-fourth of the contents of the bottle; (2) a 90-sec. interval, after which the infant was permitted to take another fourth of the formula; (3) a 90-sec. period after which the infant was removed from the bassinet and patted for air expulsion; (4) a 90-sec. period after which the infant was permitted to take the third quarter of the bottle contents; (5) a 90-sec. interval, after which the last quarter of the contents of the bottle were offered; and (6) a final period of 90-sec. observation, after which the infant was returned to the hospital nursery. One departure was made from the above procedure. If during any of the 90-sec. periods of observation, the infant cried for 5 continuous seconds, the bottle was offered immediately.

During each of the between-feeding intervals, in addition to the automatic record of general body activity made on the moving paper strip of a polygraph, a record was made by the experimenter on the same paper tape of the following forms of activity: the onset and duration of any mouth activity (sucking, lip pursing, mouth-opening or closing, mouthing or sucking of hands or clothing), the onset of crying, and any other relevant type of behavior not recorded automatically. A time-marker registered five-second intervals. An experimental record was thus obtained of the latency and duration of mouth-activity and general body activity, and of the latency of crying after each fourth of the formula had been consumed.

The above procedure, except for the exact timing of the removal of the bottle, is fairly representative of the customary method of bottle-feeding, and the reactions which occurred were probably those of the usual bottle-fed infant.

Five of the 7 infants were brought into the hospital at the age of three months in an effort to discover whether the individual differences noted at birth had persisted to the age of three months.

Results

Reactions of the infants to withdrawal of the bottle after each fourth of the formula had been taken are expressed in the following ways:

(1) Amount (*i.e.*, *duration*) of general activity and of mouth activity in the interval following removal of the bottle. These amounts are computed as the *percent* of five-second periods containing the specified reaction within the interval, since the post-feeding intervals were terminated before the end of 90 sec. if crying occurred. In order to avoid spurious results due to too-short periods of observation, the period was discarded in computation of amounts of bodily activity and mouth activity if crying occurred within 15 sec after removal of the bottle. (2) Interval of time (*latency*) between removal of the bottle and the onset of any of the following behavior items: general bodily activity, mouth activity (excluding crying), and crying. Latencies were computed for each subject during each of the intra-feeding periods of observation and are expressed as the median number of seconds (to the nearest 5 seconds) before the onset of the activity after removal of the bottle. Results for the group are averages of these medians.

For convenience, the 6 observational periods of each feeding are termed: (1) *Pre-feeding period*, the period before the first food intake, (2) *Bottle I*, period following ingestion of first fourth of formula, (3) *Bottle II*, period following ingestion of second fourth of formula, (4) *Bubble* period, the period after removal from crib for air expulsion, (5) *Bottle III*, following ingestion of third fourth of the formula, and (6) *Bottle IV*, following completion of feeding. During the neonatal period all of the infants appeared 'satisfied' at the end of feeding.

Group Averages

Figs. 1 to 5 [1] show the averages for the group. Fig. 1 shows that there is a steady and significant decrease in the amount of general body activity after increasing amounts of milk have been consumed. The relative decrease is proportional to the amount of milk consumed, *i.e.*, there is less difference in the amount of activity after removal of the bottle after Bottles I and II than there is between Bottles III and IV.

Fig. 2 shows the percent of five-second intervals containing mouth activity (aside from crying) during successive periods following food intake. Unlike general bodily activity, mouth reactions are almost absent in the pre-feeding period, and become prominent only after feeding has actually taken place. Quantity of mouth activity in no period approximates the quantity of general bodily activity, and the curve of decreasing frequency as feeding progresses in no case shows as steep a gradient.

Figs. 3–5 show the group averages for the latency of the various behavior items during the pre-feeding period and following the removal of the bottle after various amounts of formula have been taken. The curves for each of the three items, crying, general activity, and mouth reactions, show the same general trend of increasing latency of the reaction as more formula has been taken. The most outstanding differences in the curves for each type of reaction are present during the pre-feeding period and during the interval following the return of the infant to the crib after removal for expulsion of air. In the pre-feeding period, there is so little mouth activity that the median is 90+ sec.[2] for the group. Crying has its shortest latency in the pre-feeding period, while general body activity has a slightly greater latency of onset in the pre-feeding period than in the periods 'Bottle I' and 'Bottle II.' The period after the 'bubble' produced relatively

[1] Each curve represents the average of 10 to 14 feedings on all except two twin infants who, because of failure to gain properly, were observed for 6 feedings each, distributed to include the third, fifth, sixth and tenth days of their hospitalization.

[2] A numerical value of 90+ signifies that the behavior item did not occur within the 90 – sec. period of observation.

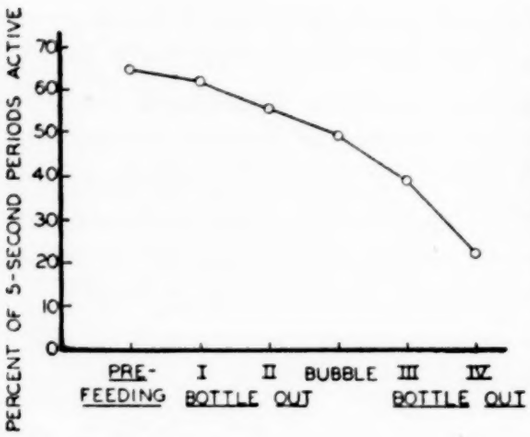


FIG. 1. Amount of general bodily activity after successive withdrawals of bottle. Average all subjects.

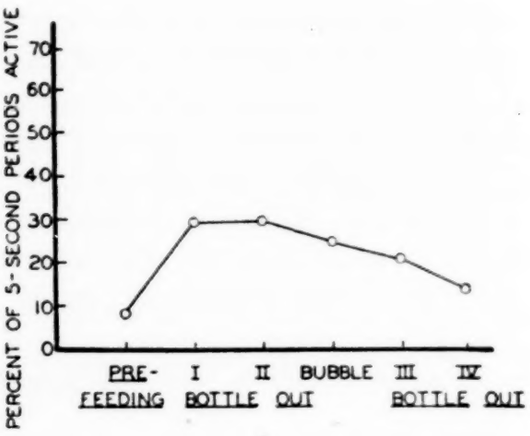


FIG. 2. Amount of mouth activity after successive withdrawals of bottle. Average all subjects.

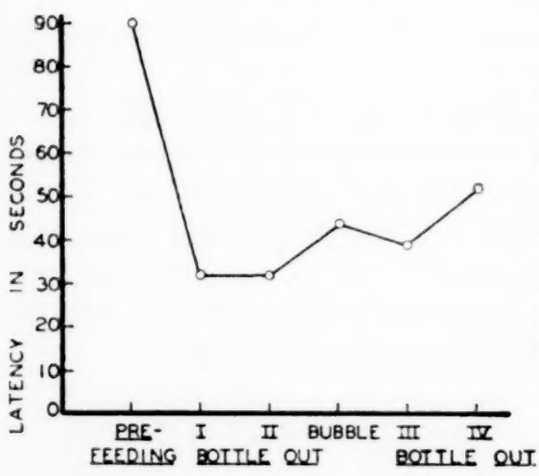


FIG. 3. Latency of mouth reactions after successive withdrawals of bottle. Average of medians of all subjects.

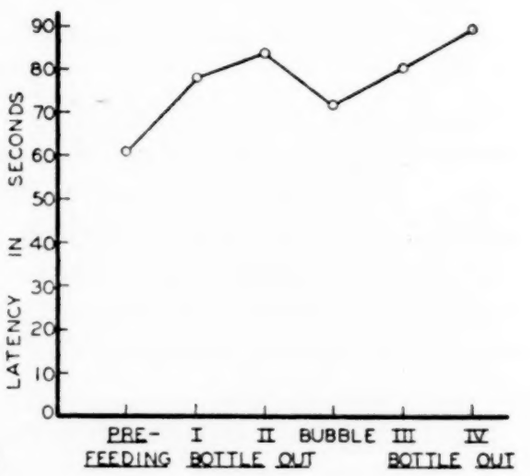


FIG. 4. Latency of crying after successive withdrawals of bottle. Average of medians of all subjects.

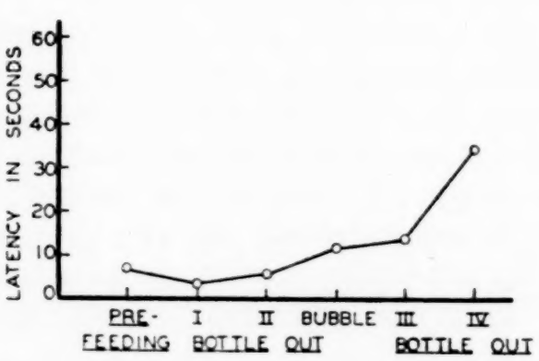


FIG. 5. Latency of general bodily activity after successive withdrawals of bottle. Average of medians of all subjects.

greater latencies for general activity and mouth activity but a shorter latency for crying.

Figs. 6–10 show curves of reaction for individual subjects. The solid lines represent results at the newborn stage, the broken lines the averages for two feedings at the age of three months. The curves for the newborn period show that although there was present in all subjects the trend towards decrease in reaction to withdrawal of the bottle as satiation increased, both the quantitative and qualitative aspects of the reactions showed considerable individual variation. Table I presents data concerning individual subjects, relevant to interpretation of the individual curves. The Don twins who were

TABLE I

Subject	Sex	Birthweight (Grams)	Weight on Discharge from Hospital	Pediatrician's Discharge Diagnosis	Ave. cc. per Feeding (during Period of Experiment)
Con	F	2865	2815	Normal newborn	50
Dec	M	3750	3820	Normal newborn	65
Don	F twins	2250	2150	Discharged against advice; premature weight	32
Don	M twins	2305	2315	Discharged against advice; premature weight	41
Gar	F	4310	4485	Normal newborn	64
Han	F	3010	3095	Normal newborn	63
Tay	F	3655	3875	Normal newborn	60

smallest and who were discharged at the end of ten days in the hospital 'against advice' because of birthweight under five pounds, showed a generally low level of response in the experimental situation. They never appeared 'hungry' and often refused part of the contents of the bottle. Their reactions were limited chiefly to a general restlessness; they never cried during the experimental period and showed very little mouth activity. Dec, on the other hand, gave strong responses of each type, showing five times as much general bodily activity to withdrawal of the bottle as did Don Boy, the least active of all subjects.

Table II shows that the rank order inter-correlations for the various reactions were high.

TABLE II

	Rank Order Correlation
Amount of mouth activity and amount of general bodily activity	.86
Amount of general activity and frequency of crying	.96
Amount of mouth activity and frequency of crying	.92
Amount of general activity and birthweight	.72
Latency of general activity and amount of general activity	−.81
Latency of mouth activity and amount of mouth activity	−1.00

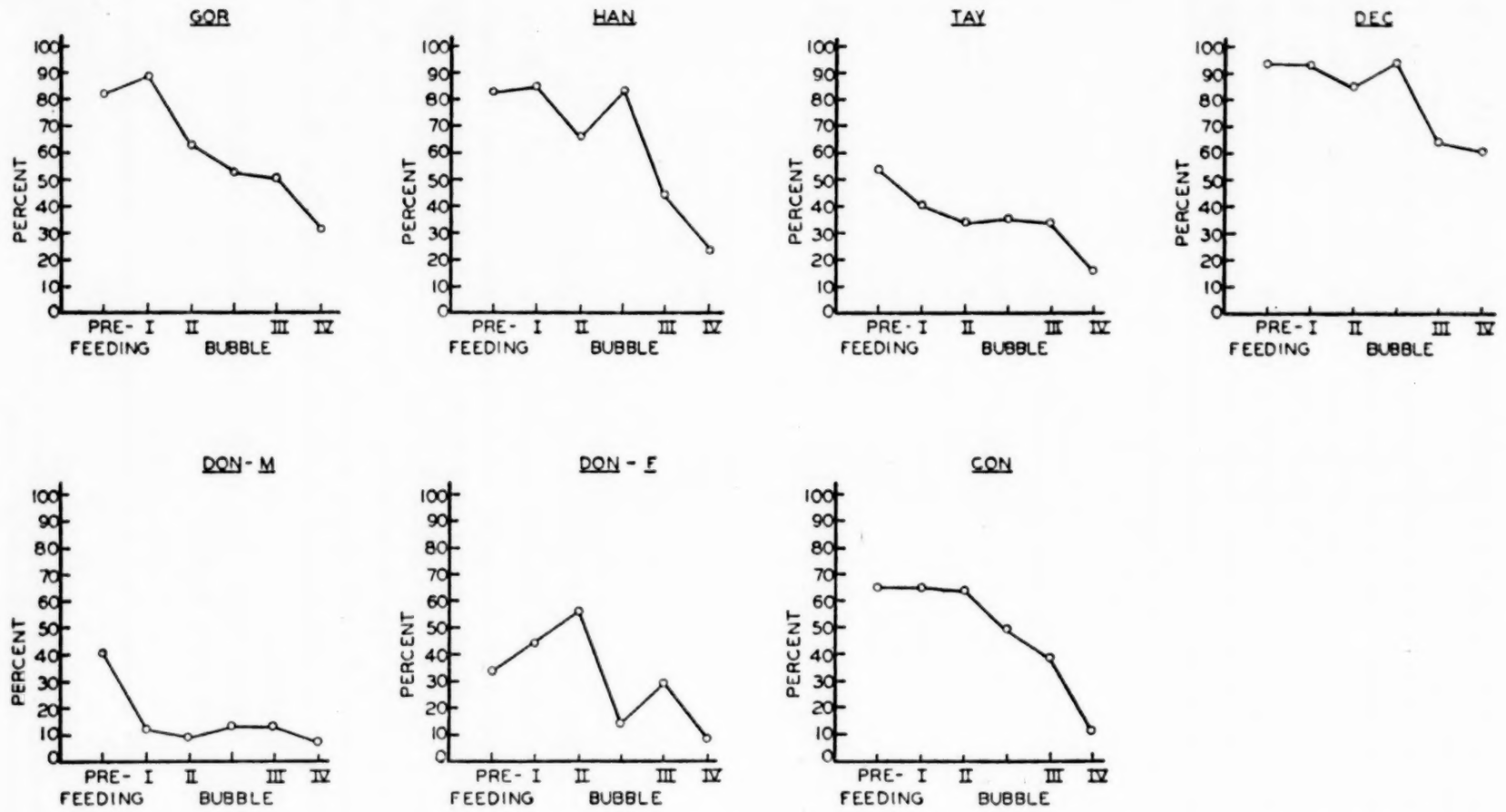


FIG. 6. Duration of general bodily activity. Curves for individual subjects. Percent of time active in successive interfeeding intervals. I, II, III, and IV refer to successive withdrawals of bottle.

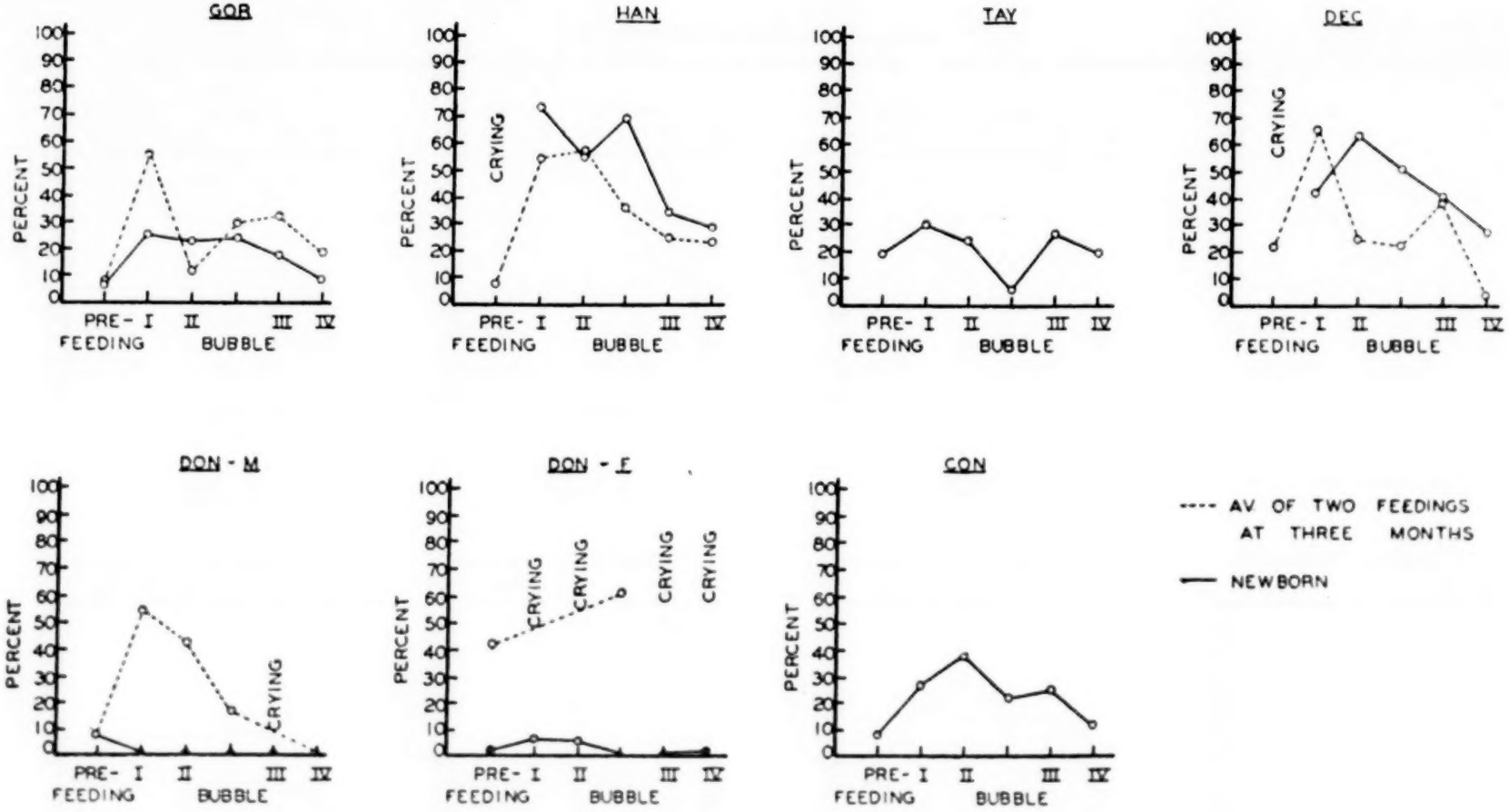


FIG. 7. Duration of mouth activity. Curves for individual subjects. Percent of time containing mouth activity in successive inter-feeding intervals. I, II, III, and IV refer to successive withdrawals of bottle. Comparison of responses at the newborn and three-months period of observation.

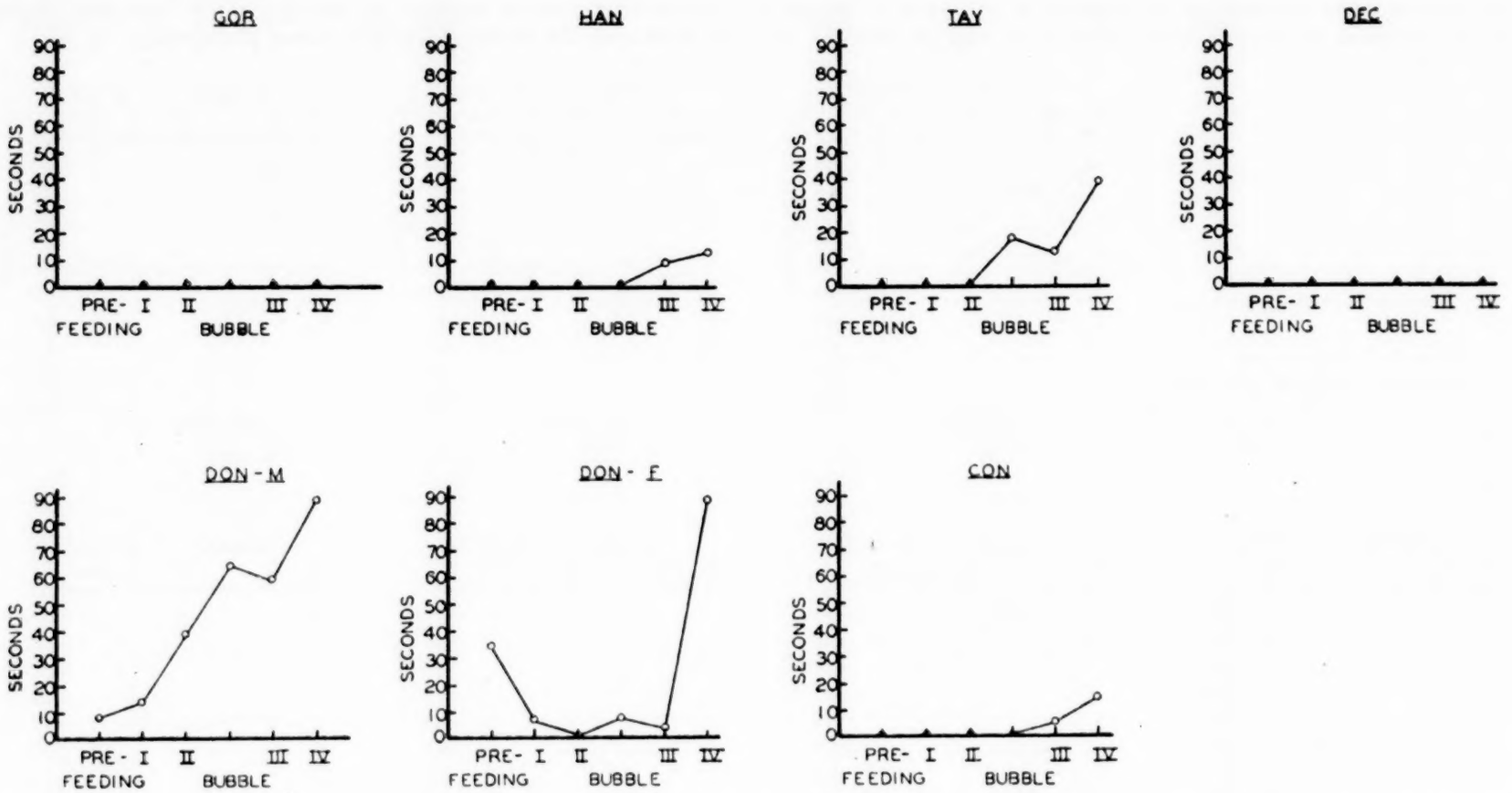


FIG. 8. Median latency of general bodily activity in successive interfeeding intervals. Curves for individual subjects. I, II, III, and IV refer to successive withdrawals of bottle.

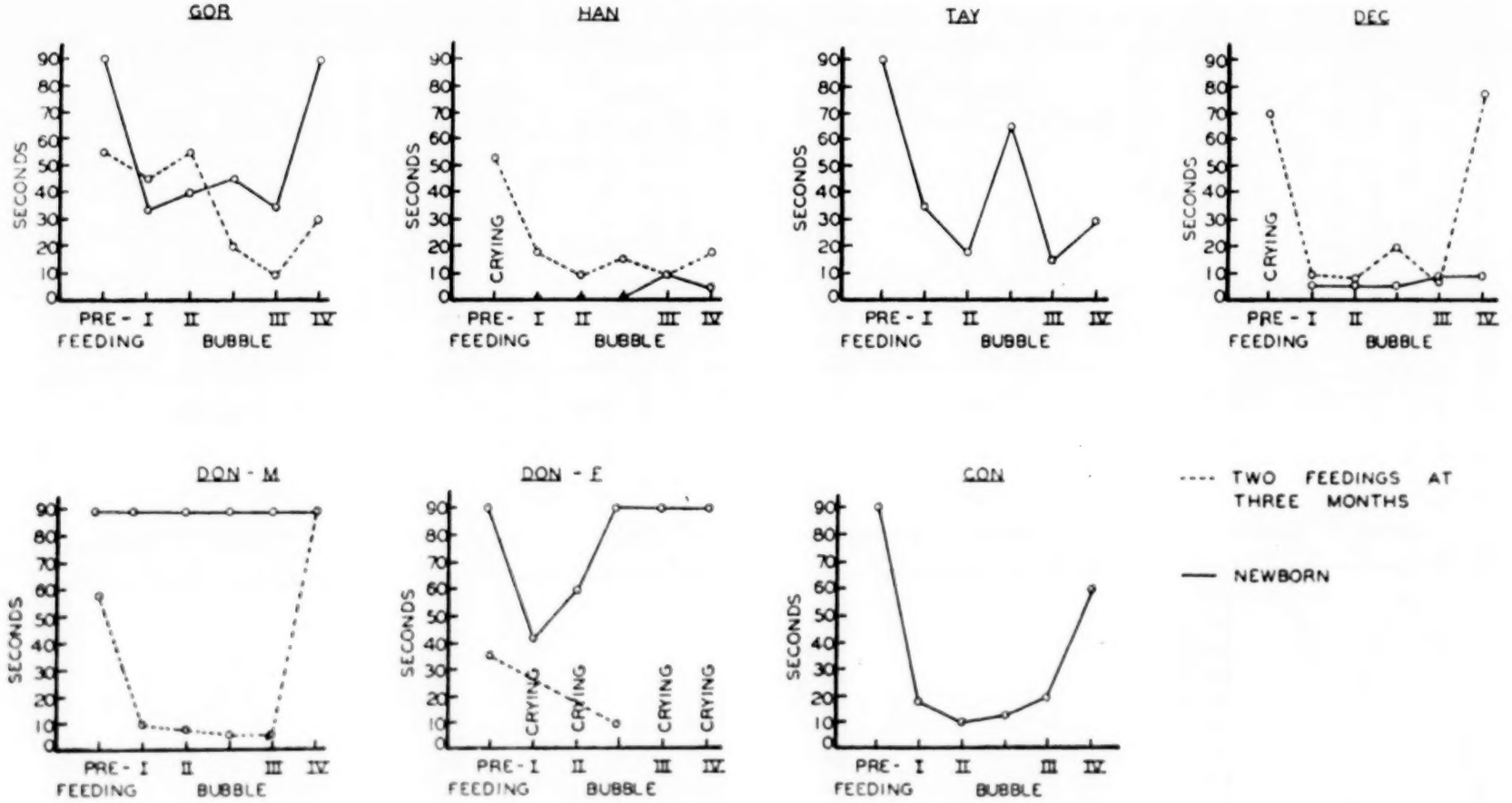


FIG. 9. Median latency of mouth reactions in successive interfeeding intervals. Curves for individual subjects. I, II, III, and IV refer to successive withdrawals of bottle. Comparison of newborn and three-months periods of observation.

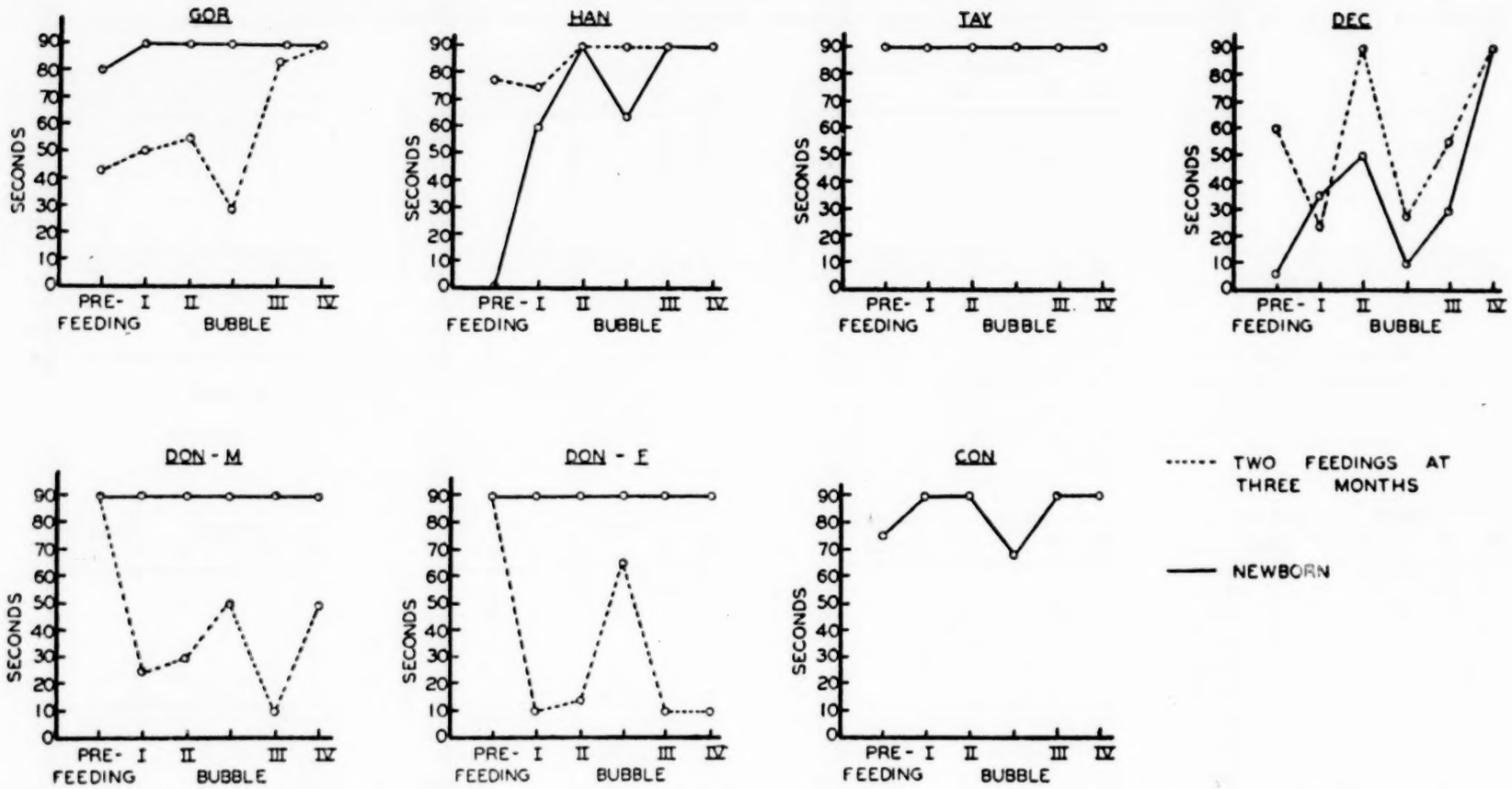


FIG. 10. Median latency of crying in successive interfeeding intervals. Curves for individual subjects. I, II, III, and IV refer to successive withdrawals of bottle. Comparison of newborn and three-months periods of observation.

If results for the atypical Don twins are omitted, because of their general lack of response, there is little correlation between weight and activity level, and between mouth activity and general level of activity. The correlation between crying and general activity however was in all subjects high.

Table III presents results for individual subjects, comparing the

TABLE III

	Average Percent General Bodily Activity		Average Percent Mouth Activity		Latencies. Av. of Medians (Seconds)					
					General Activity		Crying		Mouth Activity	
	B1 & 2	B3 & 4	B1 & 2	B3 & 4	B1 & 2	B3 & 4	B1 & 2	B3 & 4	B1 & 2	B3 & 4
Con	65	25	33	19	0	10	90+	90+	14	40
Dec	90	63	53	35	0	0	42	60	5	10
Don ♀	51	20	6	1	4	47	90+	90+	51	90+
Don ♂	11	10	0	0	12	75	90+	90+	90+	90+
Gor	76	42	23	14	0	0	90+	90+	37	63
Han	77	35	64	32	0	12	75	90	7	22
Tay	38	26	27	24	0	27	90+	90+	27	23

reactions to the first and second withdrawal of the bottle with those to the third and fourth withdrawal. There is an average of 42 percent decrease in the amount of general activity in the periods following the third and fourth withdrawals, and a decrease of 36 percent in amount of mouth activity. Table IV shows the rank order correlations of

TABLE IV

RANK ORDER CORRELATIONS OF REACTIONS TO WITHDRAWAL OF BOTTLES I AND II WITH THAT OF WITHDRAWAL OF BOTTLES III AND IV

Percent general bodily activity	.90
Percent mouth activity	.93
Latency general bodily activity	.81
Latency mouth activity	.96

reactions to withdrawal of Bottles I and II with those to withdrawal of Bottles III and IV. Analysis of Tables II, III, and IV and of Figs. 6–9 shows that there were not only individual quantitative differences in the frustration-reaction pattern, but also qualitative differences. Mouth activity, if the atypical Don twins are omitted, varied independently of body weight and of general activity level. The high correlations of Table IV show that there is consistency of reaction-pattern under different degrees of frustration. That is, the infants showing the most general activity and mouth activity at the height of frustration continued to show a higher proportion of these reactions as the degree of frustration lessened. A similar trend was

also shown with latency of the reactions. Interesting is the fact that although the rank order correlation of the amount of general activity with the latency of general activity is only − .81, the correlation of the amount and latency of mouth reaction is − 1.00.

Frustration-Reactions at the End of Three Months

Five of the 7 infants were brought back to the hospital for observation of two feedings at the end of three months. It was discovered that three of the five (Don twins and Gor) at this time were feeding problems in the sense of not being satisfied at the end of feeding. All infants at this time were receiving 5 ounces of milk per feeding and had not yet been given cereal or other solid foods. In the case of the Don twins, the 5 ounces consisted of half-skimmed milk and this was supplemented by two ounces of boiled water.

The same experimental procedure was carried out as formerly.

The broken lines on Figs. 7, 9, and 10 give the results for individual subjects at the age of three months. Because of the feeding difficulties described above, and the few cases employed, the data are of little value in the study of the persistence of individual differences. A few items of interest, however, do appear. It was found that the duration and latency of general bodily activity could no longer be used as a differential criterion of frustration since all infants were active throughout the feeding period. At the three-months test, three infants, Gor and the Don twins, were by definition more 'frustrated' in the feeding situation than during the neonatal test periods. They were apparently receiving too little food and were consequently dissatisfied at the end of feeding. In these infants it was found that the latency of crying and the latency of mouth activity were decidedly shorter, and the amount of mouth activity greater than during the neonatal period. With the other two infants, Dec and Han, who appeared satisfied after feeding, there was a slight increase in the latency of crying and mouth activity and a decrease in the amount of mouth activity. Infant Han, however, who had ranked highest in mouth activity during the neonatal period, still showed a relatively high proportion of mouth reactions in the frustration situation, in spite of the fact that she was one of the two infants who appeared to have no feeding problems at this time.

The shape of the 'frustration' curves also varied somewhat from those obtained during the neonatal period. The increase in mouth activity after the pre-feeding period was more marked at three months, and the latency of mouth activity tended to increase sharply only after the last fourth of the formula had been taken.

Summary of Results

The above data may be summarized in the following statements: 1. Newborn infants show marked frustration reactions in the feeding situation, the amount of reaction varying with the strength of instigation of the hunger drive. Decreasing degrees of the hunger drive were accompanied by decreasing amounts of general activity and mouth activity, and by increasing latencies of general activity, crying, and mouth activity after the bottle was withdrawn at the various stages of feeding. 2. The pattern of the frustration-reaction in newborns appeared to have fairly well-defined characteristics. In the pre-feeding period, the infants were generally restless and frequently cried. As has been shown in a previous study (**6**) the four-hour wait between feedings is probably in itself a frustrating situation for the newborn and one which leads to random activity. Once feeding has begun, however, the hunger goal appears to become more defined, and the infants respond to blocking of feeding by immediate general activity and by goal-directed mouth activity. In other words, the frustration becomes greater as the feeding pattern becomes more specific. 3. There are individual differences in the manner in which newborn infants react to frustration in the feeding situation. This was manifested chiefly in two ways, in the degree of general reactivity, and in the qualitative nature of the response. For example, the infant showing most reaction to withdrawal of the bottle (Dec) showed 5 times as much general bodily activity in the frustrating situation as did the least reactive infant (Don ♂). The intercorrelations between amount of crying, amount of general activity and birthweights were high when results for all subjects are included. This probably means that the smaller infants were less 'hungry.' However, a qualitative difference in the frustration-reaction pattern of individual infants was shown in the amount of mouth activity occurring in the inter-feeding intervals. Some infants, regardless of their general activity level and their birthweights, showed more of the oral component in their reactions than others, and this reaction-pattern persisted even after some degree of satiation was approached. 4. There were several differences in the reaction-pattern to frustration at the age of three months. It was no longer possible to employ the latency or degree of bodily activity after removal of the bottle as a criterion of frustration since all infants were in constant activity throughout the period. The infants who were receiving an adequate diet showed somewhat less frustration-reaction than they had shown at birth, whereas those infants who were 'dissatisfied' with their diets, showed an increased intensity of reaction. The Don twins, who at birth never appeared hungry and almost always refused part

of the formula, and who had shown fewest reactions in the frustration situation, at three months, under the instigation of a strong hunger drive, showed more crying than any other of the infants.

One other difference which was evident at three months was the difference in the shape of the curves for mouth reactions. There is a sharper gradient between Bottles III and IV both in latency and amount of mouth activity at three months than during the neonatal period. This may be a sign of learning in the sense of the infants' having learned to have anticipated the end of feeding.

Discussion

From the methodological point of view, the results of the present experiment suggest that the technique of frustration in the feeding situation is perhaps a more effective way of studying emotions in the neonate than the traditional methods of Watson and Morgan (10). Irwin (3, 4), Sherman and Sherman (8), Taylor (9), Dennis (1), and others have shown that the emotional reactions described by Watson and Morgan are not typically elicited in the newborn infant. Certainly frustration in the feeding situation is a more common experience of the infant in this culture. Very few newborn infants are subjected to frequent sudden loud noises, loss of support, or repeated restriction of the nose or extremities. However, it is a biological necessity that they make adjustment to the frustrations which occur all too frequently in the feeding situation. The question as to whether the reactions shown in this experiment can be termed 'emotional reactions' may be debatable, but is, I believe, defensible. Although no one has satisfactorily defined emotional behavior, one of the most commonly recognized emotional states is that form of disorganized behavior which results when an organism is prevented from making a goal response. No observer, looking at a newborn infant who has not been fed for four hours, could avoid the assumption that widespread internal changes resulting from the hunger drive are producing concomitant widespread effects upon the infant's organs of overt response. In the above experimental situation, the degree and quality of the 'emotional' reactions appear to be dependent on two factors: the degree of frustration, and the reaction characteristics of the individual infant. The data prove that the stronger the hunger drive (as measured by reactions during the first stages of feeding) the greater is the reaction to feeding frustration. In regard to the second point, concerning the qualitative nature of the emotional response in individual cases, the results of the study suggest that there are innate characteristics existing from birth which determine an infant's direction of response in a frustrating situation. Certain infants have a

generally higher level of response even when the hunger drive approaches satiation. Similarly, one infant, both during the neonatal period and at three months, showed a tendency toward a relatively higher percentage of mouth reactions than other infants. The last point is consistent with Freud's concept of innate differences in the oral libidinal instinct.

The general pattern of reaction in the feeding situation in the neonatal stage and the change in pattern of general bodily activity at the three months level lends support to Kleitman's theory (**5**) of the development of the 'wakefulness of choice' from the 'wakefulness of necessity.' Kleitman states that the newborn is awake only when he is hungry, wet, cold, or otherwise disturbed, falling asleep again after the disturbance is removed. 'Wakefulness of choice' develops as the cerebral cortex develops greater functional activity. The subjects of this experiment in the neonatal period tended to become more and more quiescent as hunger was satiated; whereas at three months the level of general activity (aside from crying) remained fairly constant throughout the feeding situation, and could no longer be employed as a differential criterion of frustration. This probably means that by virtue of increasing cortical differentiation, or at any rate by means of a process of learning and acculturation, the response of general bodily activity has become attached to an increasing number of stimuli (*e.g.*, sight of the mother, kinesthetic cues of the feeding posture, internal cues relating to satiation, etc.).

The distribution of mouth activity in the frustration situation also brings up several interesting facts for discussion. The fact that mouth activity begins in most cases only after feeding has taken place and then occurs with decreasing frequency as feeding progresses can be explained in terms of two concepts: (1) the perseveration of a neuromuscular set initiated by the act of feeding, and (2) the concept of the hunger drive which, as it diminishes in strength, no longer produces internal stimuli and resultant strong responses, oral or otherwise. The individual differences in amount of mouth activity, however, are probably due to inherent differences in neurophysiological structure.

Summary

Frustration in the feeding situation was investigated in 7 full term newborn infants by studying the latency and the amount of crying, general bodily activity, and mouth activity after each fourth of the contents of the bottle had been consumed. Both the latency and the amount of these reactions varied consistently with the amount of milk which had been taken. It was concluded that newborns show decided frustration-reactions, the strength of the reaction varying

with the strength of instigation of the hunger drive. Although all infants showed decreasing amounts and increasing latencies of reaction as feeding progressed, individual differences both in quantity and quality of the reactions were shown.

Although it seems probable that the infants were generally 'frustrated' before feeding began, by virtue of having waited four hours for feeding, the frustration-reaction pattern changed somewhat once feeding had begun. Although general activity and crying were prominent in the pre-feeding control period, mouth activity became prominent only after feeding had begun. This was explained in terms of the persistence of the neuromuscular set initiated by the act of feeding, and was found to be present in all infants.

When 5 of the infants were returned for observation at the age of three months, the presence of feeding problems in 3 of the 5 obscured results to some extent. Some of the observations however were important at this stage. The pattern of the frustration-reaction was similar to that seen at birth, except that all the infants showed general bodily activity throughout the feeding period and did not relapse into a quiescent state at the end of feeding as they had done at the neonatal stage. This was discussed in relation to two points: first, as marking the earliest stages of the development of Kleitman's 'wakefulness of choice,' and second, as showing indications of adaptation or learning in the feeding situation. One other observation at the three-months' period was the persistence of individual differences in general activity level and in the relative proportion of mouth activity. The latter was cited as evidence for Freud's theory of individual differences in the oral libidinal instinct.

(Manuscript received September 21, 1942)

References

1. Dennis, W. Infant reaction to restraint: an evaluation of Watson's theory. *Trans. N. Y. Acad. Sci.*, 1940, **2**, 202–218.
2. Dollard, J., Doob, L. W., Miller, N. E., Mowrer, O. H., & Sears, R. R. *Frustration and aggression.* New Haven: Yale Univ. Press, 1939.
3. Irwin, O. C. Infant responses to vertical movement. *Child Develpm.*, 1932, **3**, 167–169.
4. Irwin, O. C. The latent time of the body startle in infants. *Child Develpm.*, 1932, **3**, 104–107.
5. Kleitman, N. *Sleep and wakefulness.* Chicago: Univ. Chicago Press, 1939.
6. Marquis, D. P. Learning in the neonate: the modification of behavior under three feeding schedules. *J. Exp. Psychol.*, 1941, **29**, 263–281.
7. Sears, R. R., & Sears, P. S. Minor studies of aggression: V. Strength of frustration-reaction as a function of strength of drive. *J. Psychol.*, 1940, **9**, 297–300.
8. Sherman, M., & Sherman, I. C. *The process of human behavior.* N. Y.: Norton, 1929.
9. Taylor, J. H. Innate emotional responses in infants. *Ohio State Univ. Stud.*, 1934, No. 12, 69–81.
10. Watson, J. B., & Morgan, J. J. B. Emotional reactions and psychological experimentation *Amer. J. Psychol.*, 1917, **28**, 163–174.

THE PSEUDO-CONDITIONED EYELID RESPONSE [1]

BY DAVID A. GRANT

University of Wisconsin

INTRODUCTION

Pseudo-conditioning, like orthodox conditioning, results in the evocation of responses by previously inadequate stimuli. Pseudo-conditioning is brought about by the repetition of the unconditioned stimulus *alone;* unlike orthodox conditioning, there is *no pairing* of the conditioned and unconditioned stimuli during the training series. Although much earlier work bears on the problem, the importance of the phenomenon and its relation to conditioning has been recognized chiefly since Grether (**6**) called attention to it.

A number of previous studies on the eyelid response have included observations which have some bearing upon pseudo-conditioning or have included control procedures designed specifically to check on the extent of sensitization resulting from pseudo-conditioning reinforcement.

Peak (**23**) observed secondary eyelid responses in some of her subjects following repetitive auditory stimulation. She remarked (**23**, p. 20) that the phenomenon ". . . bears the look of conditioning, and would seem to be both a conditioned voluntary and conditioned reflex, since it has certain of the characteristics of the voluntary and conditioned reflex!"

Hilgard (**12**), in experiments with human subjects, and Hilgard and Marquis (**14, 15, 20, 21**), in experiments with dogs and monkeys, introduced limited pseudo-conditioning control procedures before conventional conditioning reinforcement was given. No pseudo-conditioning was observed. Under the conditions of these experiments, however, it took at least *two to four* days for extensive conditioning to appear, and the control procedures were used *for one day only*. This indicates that as far as pseudo-conditioning was concerned, the control procedures were inadequate. Had these procedures been tried for two to four days, they might have yielded positive results. An additional factor which very likely decreased pseudo-conditioning in the animal studies was an adaptation series of 60 conditioned stimuli on the day before the pseudo-conditioning trials.

[1] This experiment was carried out in the psychological laboratory of Stanford University; an abstract has appeared (**4**). The writer wishes to thank Professor E. R. Hilgard for advice and suggestions at all stages of the study.

Bernstein (2) reported that the average amplitude of the eyelid *reflex* to sound was augmented in four subjects following a series of isolated shocks to the cheek. This facilitative tendency, to the extent that it was experimentally demonstrated, can be considered pseudo-conditioning, but the number of subjects was very small, and the nature (reflex or non-reflex) of the facilitated responses was not clearly shown. The disparity between Bernstein's findings and those of Hilgard and Marquis were most likely caused by differences in experimental arrangements.

In view of the conflicts in the scant available data it has not been possible to estimate the extent to which the eyelid response is subject to pseudo-conditioning. The present study was designed to clarify further the nature of eyelid pseudo-conditioning and the extent to which pseudo-conditioning operates in the eyelid response when stimuli and experimental conditions favor the appearance of sensitization.

Apparatus and Procedure

The stimulating and recording apparatus were similar to that of Hilgard and Campbell (13). Eyelid reactions were photographically recorded from the movements of a lever system connected to the eyelid by a thread and a small piece of adhesive tape. The lever system was devised by Lumsdaine (19), and gave linear amplification with relatively little distortion due to mechanical factors.

The *unconditioned stimulus* was a puff of air to the left cornea. The four *conditioned stimuli* (actually, *pseudo-conditioned stimuli*) consisted of a light stimulus and three different intensities of auditory stimulation from an ordinary signal buzzer (designated B_1, B_2, and B_3).

The light stimulus consisted of the sudden increase in the brightness of a 10 cm square milk-glass screen from a pre-stimulus level of about .02 millilamberts to about 300 millilamberts.[2] The screen was situated 60 cm directly in front of the subject. The light stimulus lasted for about 550 msec. The subject was instructed to fixate visually at all times a small black cross which was visible in the center of the milk-glass screen.

The three buzzer stimuli varied in frequency and intensity. They were spaced at approximately equal sense-distances on the basis of the averaged judgments of four subjects. The buzzer stimuli lasted approximately 650 msec.

Instructions designed to obtain a relatively non-inhibitory set—*cf.* Miller (22) and Hilgard and Marquis (16, pp. 263–265)—were read to the subjects at the beginning of each experimental session.

A total of 61 students from sophomore classes in psychology served as subjects. Of these, 31 were used in the experimental group, and 30 were used in the control group.

On the first day of the experiment all subjects were first given a *pre-test* consisting of two trials with each of the four conditioned stimuli.[3] The purpose of the pre-test was to record the original responses to the conditioned stimuli.

For the *experimental group* the pre-test was followed by a series of 40 puff stimuli *not paired with any of the conditioned stimuli.* Following the 40 'puff-alone' trials, a post-test was given, consisting of two trials with each of the four conditioned stimuli in balanced order. Comparison of the responses on the pre- and post-tests permitted evaluation of the results of the intervening experimental procedure.

The *control group* was given the pre-test and the post-test, but instead of receiving puff stimulation, members of this group simply sat and maintained their fixation on the fixation cross.

---

[2] Photometric measurements were made with a Macbeth Illuminometer.

[3] The conditioned stimuli were always presented in balanced permuted cycles (*e.g.*, L, B_1, B_2, B_3, B_4, B_2, B_1, L; and B_1, L, B_3, B_2, B_2, B_3, L, B_1; etc.), and the 24 possible permutations of these cycles were systematically varied throughout the experiment.

No stimuli were presented, but a time interval elapsed equal to that devoted to the total puff series in the experimental group. *The control group was used for one day only.*

The members of the experimental group returned the next day, and the procedure was the same as the experimental procedure of the first day. Then, following a retention period of at least one week,[4] members of the experimental group were tested for retention and extinction of their pseudo-conditioned eyelid responses. The test for retention and extinction consisted of six stimulations with each of the four conditioned stimuli (three successive permuted cycles; *cf.* footnote 3). No puff stimulation was given during the retention and extinction tests on the third day.

Results

1. *Reactions to light.*—Following both the experimental and control procedures, striking and consistent changes were found in the eyelid reactions to the light stimulus. Increases in the frequency of eyelid reactions appeared, and these increases were similar to the increases obtained in orthodox conditioning experiments.

a. Latency of sensitized responses to light.—A problem which immediately arose was whether or not the frequency increases involved sensitization at the reflex level, or whether the responses showing frequency increment had the latency characteristics of ordinary conditioned responses. Systematic consideration of the latencies of all eyelid responses helped to settle this issue. In Fig. 1 is shown the distribution of latencies of all eyelid closures or partial closures recorded from members of the pseudo-conditioning group

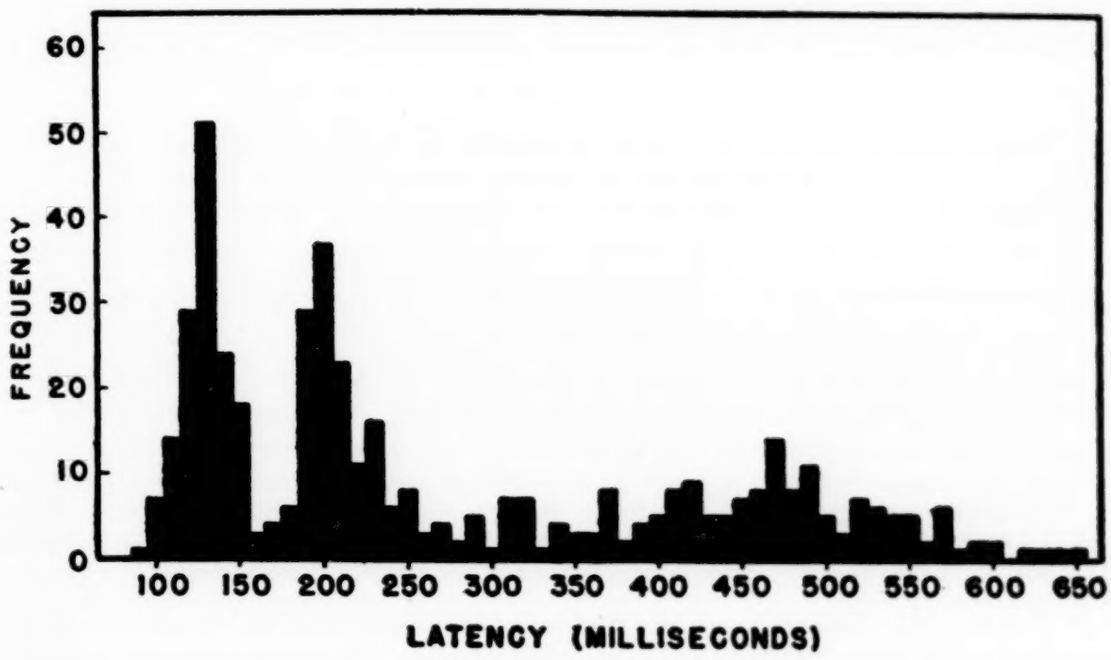


Fig. 1. Distribution of the latencies of all eyelid responses evoked by the light stimulus in the pseudo-conditioning group. The ordinate shows the frequency of responses at each latency. Latencies are as recorded and are not corrected for the constant latency of recording devices (*cf.* footnote 5).

[4] Although the minimum retention period was one week, most of the subjects returned on the ninth or tenth day. One failed to return.

during the presentation of the light stimulus. The frequency of responses is plotted on the ordinate axis against latency in msec. along the abscissa.

The latency distribution in Fig. 1 is distinctly bimodal. The first mode, at 130 msec.,[5] consists chiefly of reflexes to the light stimulus. If pseudo-conditioned eyelid responses consisted merely of reflex sensitization, these would be the reactions which would increase in frequency of appearance. However, the reflex responses typically decreased in frequency and amplitude during the experiment, thus indicating that sensitization, if it occurred, was obscured by the usual course of adaptation.

There is, however, a definite mode of responses at 200 msec. which lies within the usual latency range of conditioned eyelid reactions (*i.e.*, between 170 and 450 msec.). These responses were the ones which showed clear-cut increases in frequency of appearance in both the experimental and control groups. Further discussion of the eyelid responses observed in this study will be limited to those with latencies between 170 and 450 msec.[6]

b. Changes in frequency.—One of the most common measures of conditioning is the degree to which the responses designated as conditioned increase in frequency during the training period. That such increases appear within a similar period of pseudo-conditioning training without *pairing* of conditioned and unconditioned stimuli is shown in Table 1. In Table 1 are given the percent frequencies of

TABLE 1

PERCENTAGE FREQUENCY OF EYELID RESPONSES TO THE LIGHT STIMULUS WITH LATENCIES OF 170 TO 450 MILLISECONDS

Group	Day 1		Day 2		Day 3		
	Pre-test	Post-test	Pre-test	Post-test	1	2	3
Pseudo-conditioning ($N = 31$)	19	58	27	71	42	60	67
Fixation-control ($N = 30$)	25	53					

Each percentage frequency is based upon records of eyelid reactions to two stimulations per subject with the light during a balanced permuted cycle of conditioned stimuli. There were 31 subjects in the pseudo-conditioning group and 30 in the control group. One subject in the pseudo-conditioning group failed to appear for the third day of the experiment.

[5] This and other latencies herein reported are as determined from the records and are not corrected for the constant latency of the recording devices. The absolute latency is in each instance approximately 50 msec. less than that reported.

[6] Conditioned eyelid responses to light typically have measurable latencies up to about 450 msec. The unconditioned stimulus is presented about 400 msec. after the onset of the light, and conditioned responses with latencies greater than about 450 msec. are obscured by reflexes to the unconditioned stimulus. Choice of these latency limits permits comparison of the present findings on pseudo-conditioning with the findings in previous conventional eyelid conditioning experiments.

reactions evoked by the light stimulus from members of the pseudo-conditioning and control groups during the pre- and post-tests of days 1 and 2 and during the extinction tests of day 3. The day 3 data are grouped into three successive 8-trial cycles. The reliability of frequency changes from test to test is indicated by the critical ratios of the differences which appear in Table 2. Negative critical ratios refer to frequency decreases.

TABLE 2

CRITICAL RATIOS OF CHANGES IN FREQUENCY OF PSEUDO-CONDITIONED AND CONTROL GROUP RESPONSES

Group	Increase Within Day 1	Decrease Overnight Between Days 1 and 2	Increase Within Day 2	Decrease Overnight Between Days 2 and 3	Day 3 Increases	
					Cycles 1 to 2	Cycles 1 to 3
Pseudo-conditioning	4.96	−3.51	6.12	−4.08	2.84	3.81
Fixation control	3.25					

In Table 1 the frequency of responses for the pseudo-conditioning group increases markedly from 19 percent on the pre-test of day 1 to 58 percent during the post-test following the 40 puff trials of the first day. The critical ratio for the increase within day 1 was 4.96, indicating that great confidence may be placed in its reliability. This constitutes highly reliable evidence for pseudo-conditioning. The rôle of the air-puff is not clear, however, for at the same time the control group, with no puff stimulation, shows a frequency increase from 25 percent on the pre-test to 53 percent on the post-test. This increase also was statistically significant (critical ratio = 3.25). The difference between the frequency increase of the pseudo-conditioning group and the frequency increase of the fixation group was not statistically significant (critical ratio = 1.23). Thus, the repetitive puff stimulation may not be the only factor accounting for the frequency changes in the pseudo-conditioning group.

Following the first day's experiment, there is a statistically significant overnight decrease in response frequency in the pseudo-conditioning group between the post-test of day 1 and the pre-test of day 2. This is followed by a highly reliable increase on day 2. Then, during the 7–10 day retention period, there is a significant decrease in response frequency as shown by the critical ratio of 4.08 for the difference between the response frequency on the post-test of day 2 and the response frequency on the first stimulation cycle on day 3.

In spite of the overnight drops in frequency described above, there is evidence for retention of pseudo-conditioning. The decreases in frequency between days 1 and 2 and between days 2 and 3 are in

neither case great enough to reduce the response frequency to the initial (pre-test of day 1) level. The amount of overnight retention, as measured by the difference between the percent frequencies on the pre-tests of days 1 and 2, is not reliable (critical ratio = 1.33). However, following the second day of reinforcement, the retention over the 7–10 day period intervening before day 3 is statistically significant (critical ratios for the differences between the frequency on the pre-test of day 1 and the frequency on the first, second, and third stimulation cycles of day 3 are 2.81, 4.64, and 6.28, respectively).

There is no evidence for extinction on the third day of the experiment. In fact, the trend shown by the data is definitely upward, even though no puff stimulation was given. Furthermore, the total increase in response frequency from 42 percent to 67 percent during the day 3 cycles is statistically significant (critical ratio = 3.81).

c. Amplitude results.—The average amplitude of the responses to light varied from about 3 mm to about 8 mm on the different test series. The direction of the amplitude trends corresponded well with that of the frequency trends, but for the most part the amplitude increases and decreases did not approach reliability.

2. *Reactions to sound.*—In general, the results with the buzzer stimuli showed the same trends as those with the light stimuli. However, the trends of the responses to the buzzer stimuli were neither as regular nor as reliable as those to the light stimulus.

When the latencies of all eyelid responses to each buzzer stimulus were plotted, as was done in Fig. 1 with the responses to light, each distribution showed a clear-cut mode of reflex eyelid reactions to the sound. Secondary modes were not as obvious in the latency distributions of responses evoked by sound as was the secondary mode in the corresponding distribution of responses to light.

The responses elicited by the buzzer stimuli and having latencies between 170 and 450 msec. tended to increase in frequency following the 40 puff trials on days 1 and 2, and tended to decrease in frequency during the retention intervals when no reinforcement was given. In these respects the results found with the buzzer stimuli were in agreement with the results found with the light stimulus. However, in the case of the auditory stimuli the frequency trends of the control group differed from those of the pseudo-conditioning group in direction and extent. In contrast to the results with the light stimulus, visual fixation resulted in no facilitative tendency in responses to the buzzers. The frequency trends of the pseudo-conditioned responses evoked by the buzzer stimuli were irregular on the third day of the experiment, but there was no tendency toward extinction.

None of the above mentioned trends was statistically significant, but some of the critical ratios were high enough to justify some con-

fidence in the tendencies exhibited. For example, the day 1 frequency increases from pre- to post-test for responses to B_1 and B_2, the two loudest buzzer stimuli, gave critical ratios of 1.99 and 2.56, respectively, and the critical ratio for the day 2 frequency increase in responses to B_2 was 2.22. The over-all increases in frequency of responses to B_2 from 5 percent on the pre-test of day 1 to 23 percent and 31 percent on the second and third of the balanced, permuted cycles on day 3 were statistically reliable with critical ratios of 3.05 and 3.63, respectively.

Discussion

1. *The Pseudo-conditioned Eyelid Response.*—An important result of this study is the clear-cut evidence for pseudo-conditioned eyelid responses to light. The evidence consists principally of the highly reliable frequency increases observed in the eyelid responses with latencies from 170 to 450 msec. following repetitive puff stimulation. The frequency increases were retained over a period of a week and were resistant to extinction. This agrees with other observations on the stability of pseudo-conditioning reported by Sears (**24**) and Harlow and Toltzien (**8**), and is in line with Ukhtomski's criteria for a 'dominant.' [7]

The results with the visual and auditory stimuli indicate broad sensory generalization of the eyelid response tendencies. This has been found to be typical of pseudo-conditioning by investigators studying other response systems (**6, 7, 3, 17, 5**). It should be noted, however, that generalization was not complete enough to include an increase in frequency of 'incidental' winks. This was shown by negative results on control trials and on measurements of winking rate [8] during the experiment.

The observed frequency increases are perhaps of greater importance because they are similar in magnitude to frequency increases of responses within the same latency range which have been reported following paired stimulation in eyelid conditioning (**16**). And in addition to resembling conditioned responses in frequency increments, the pseudo-conditioned responses closely resemble the conditioned eyelid responses of prior studies in amplitude and latency. The resemblance between the pseudo-conditioned and conditioned eyelid responses is entirely in accord with the findings of previous studies in which both conditioning and pseudo-conditioning experiments have been reported for other action systems; *e.g.*, approach and avoidance

[7] Quoted by Ufland (**27**).

[8] Each subject's winking rate was recorded periodically throughout the experiment. In addition, control trials in which the pendulum was swung and eyelid movements recorded in the absence of stimulation were made at regular intervals in order to check for possible cues from apparatus or experimenter.

reactions of minnows and responses of frogs, amblystoma, and guinea pigs to electric shock (**25, 26**); responses of goldfish to electric shock (**7**); responses of infants to shock on the foot (**28**); the galvanic skin response in human subjects (**17**); the human startle pattern (**18**); emotional responses of monkeys (**9**) and rats (**11**); and the hand-withdrawal response (**10**). The findings of the present study go beyond the findings of the above-listed studies in that the greater accuracy of recording eyelid movements permits more detailed comparison of pseudo-conditioning and conditioning in this action system.

2. *Pseudo-conditioning and Reflex Sensitization.*—One common view appearing in the literature on pseudo-conditioning is that a pseudo-conditioned response is a sensitized reflex [9] (*cf.* **2, 24**). This certainly does not appear to be the nature of the pseudo-conditioned responses found in the present study. As reported above, the eyelid responses which showed frequency increases following the experimental and control procedures were those at latencies greater than 170 msec.[10] The primary reflexes to sound and light showed no frequency increases, indicating that if reflex sensitization did occur, it was obscured by adaptation effects. Besides differences in frequency trends, another functional difference was found between the primary reflexes and the longer-latency eyelid reactions to the buzzer stimuli. The frequency of reflex responses varied with the intensity of auditory stimulation, but the frequency of pseudo-conditioned reactions did not. Although the possibility that pseudo-conditioning consists of the sensitization of some hitherto unreported secondary eyelid reflex cannot be excluded, the present study shows clearly that the origin of the pseudo-conditioned eyelid response must be sought elsewhere than in the sensitization of the primary reflexes to sound and light.

3. *The Role of Puff Stimulation.*—One of the most noteworthy results of the present experiment is the reliable increase in frequency of eyelid reactions in the control group. Members of this group received no puff stimulation, so that the role of the air-puff in relation to the frequency increases of the pseudo-conditioning group is not clear. The implication is, of course, that the pseudo-conditioning increases were not primarily dependent upon the repetitive puff stimulation, but were caused by some other (unknown) variable in the experimental situation. If this is true, then the experimental results might better be referred to by means of some such general term as *sensitization* rather than as *pseudo-conditioning.* In any case

[9] This view may stem, in part, from Ukhtomski's concept of the 'dominant,' although the concept was extended to other than reflex responses (**1**).

[10] About 120 msec. in absolute latency. See footnote 5. Primary reflexes would almost never have latencies greater than 120 msec.

determination of the exact relationship between the puff stimulation and the present experimental results must await further research.

If the repetitive puff stimulation is not the crucial factor for the pseudo-conditioning results, then some new and important reinforcing process may have been operating. Simple explanations in terms of postural tension, darkness adaptation, suggestion, and expectancy seem inadequate to account for all the data. The crucial factor or factors remain to be identified by further experimentation.

Summary

In order to study the pseudo-conditioned eyelid response, eyelid responses were photographically recorded in an experimental group of 31 subjects and a control group of 30 subjects. Light and three intensities of sound stimulation from a signal buzzer served as four different conditioned stimuli. A puff of air to the cornea was the unconditioned stimulus.

In the experimental or pseudo-conditioning group, the procedure consisted of giving the subject 40 puff stimuli, unaccompanied by either light or buzzer stimuli, on each of the first two days of the experiment. On these two days, before and after the puff series, the subjects were given a test series consisting of two each of the four conditioned stimuli in a balanced cycle. On the third day of the experiment which followed a seven to ten day retention period, six each of the four conditioned stimuli were presented, with no puff stimulation being given on that day.

The control group was run for one day only. Members of this group were put through the same procedure as members of the experimental group on the first day, except that no puff stimuli were given. During the time intervening between the initial and final test series with the conditioned stimuli, the subject simply sat and looked at a small fixation cross. Rest periods were interspersed at corresponding times in each of the groups.

The results of the investigation may be summarized as follows:

1. Following a series of 40 puff stimuli unaccompanied by any conditioned stimuli, significant gains were found in the frequency of eyelid responses to light.

2. Although there were overnight decreases in response frequency, significant retention of the responses was shown over a period of more than a week.

3. There was no evidence for extinction of the responses during a series of 24 unreinforced presentations of the conditioned stimuli.

4. The average amplitude of responses to light followed the same trends as the frequency results, but the changes were not as reliable as those in the frequency data.

5. The reactions to the buzzer stimuli tended to follow the same trends as the reactions to light, but the changes were not as consistent nor as reliable. The intensity of the buzzer stimuli was not related to the frequency of pseudo-conditioned reactions.

6. The responses reported above resembled previously reported conditioned eyelid reactions in latency, amplitude, and frequency increment. There was no evidence for sensitization of the primary reflexes to light and sound.

7. The role of the puff stimulation in determining the increases in frequency of the reactions to light was not clear, because the fixation-control group also showed significant increases in response frequency, and members of this group received no puff stimulation.

(Manuscript received August 13, 1942)

References

1. Bekhterev, V. M. *General principles of human reflexology.* New York: International, 1932.
2. Bernstein, A. L. Temporal factors in the formation of conditioned eyelid reactions in human subjects. *J. gen. Psychol.*, 1934, **10**, 173–197.
3. Grant, D. A., & Dittmer, D. G. A tactile generalization gradient for a pseudo-conditioned response. *J. exp. Psychol.*, 1940, **26**, 404–411.
4. Grant, D. A., & Hilgard, E. R. Sensitization as a supplement to association in eyelid conditioning. *Psychol. Bull.*, 1940, **37**, 478–479.
5. Grant, D. A., & Meyer, H. I. The formation of generalized response sets during repeated electric shock stimulation. *J. gen. Psychol.*, 1941, **24**, 21–38.
6. Grether, W. F. Pseudo-conditioning without paired stimulation encountered in attempted backward conditioning. *J. comp. Psychol.*, 1938, **25**, 91–96.
7. Harlow, H. F. Forward conditioning, backward conditioning, and pseudo-conditioning in the goldfish. *J. genet. Psychol.*, 1939, **55**, 49–58.
8. Harlow, H. F., & Toltzien, F. Formation of pseudo-conditioned responses in the cat. *J. gen. Psychol.*, 1940, **23**, 367–375.
9. Harlow, H. F., & Bromer, J. A. Acquisition of new responses during inactivation of the motor, premotor, and somesthetic cortex in the monkey. *J. gen. Psychol.*, 1942, **26**, 299–313.
10. Harris, J. D. Forward conditioning, backward conditioning, pseudo-conditioning, and adaptation to the conditioned stimulus. *J. exp. Psychol.*, 1941, **28**, 491–502.
11. Harris, J. D. An analysis of certain non-associative factors inherent in avoidance conditioning in the rat. *Psychol. Bull.*, 1941, **38**, 572.
12. Hilgard, E. R. Conditioned eyelid reactions to a light stimulus based on the reflex wink to sound. *Psychol. Monogr.*, 1931, **41**, No. 184.
13. Hilgard, E. R., & Campbell, A. A. The course of acquisition and retention of conditioned eyelid responses in man. *J. exp. Psychol.*, 1936, **19**, 227–247.
14. Hilgard, E. R., & Marquis, D. G. Acquisition, extinction, and retention of conditioned lid responses to light in dogs. *J. comp. Psychol.*, 1935, **19**, 29–58.
15. Hilgard, E. R., & Marquis, D. G. Conditioned eyelid responses in monkeys, with a comparison of dog, monkey, and man. *Psychol. Monogr.*, 1936, **47**, No. 212, 186–198.
16. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* New York: Appleton-Century, 1940.
17. Hull, C. L. Explorations in the patterning of stimuli conditioned to the G.S.R. *J. exp. Psychol.*, 1940, **27**, 95–110.
18. Landis, C., & Hunt, W. A. *The startle pattern.* New York: Farrar and Rinehart, 1939.
19. Lumsdaine, A. A. Conditioned eyelid responses as mediating generalized conditioned finger reactions. *Psychol. Bull.*, 1939, **36**, 650.

20. MARQUIS, D. G., & HILGARD, E. R. Conditioned lid responses to light in dogs after removal of the visual cortex. *J. comp. Psychol.*, 1936, **22**, 157–178.
21. MARQUIS, D. G., & HILGARD, E. R. Conditioned responses to light in monkeys after removal of the occipital lobes. *Brain*, 1937, **60**, 1–12.
22. MILLER, J. The effect of facilitatory and inhibitory attitudes on eyelid conditioning. Ph.D. dissertation, Yale University. (Abstr. in *Psychol. Bull.*, 1939, **36**, 577–578.)
23. PEAK, H. Modification of the lid-reflex by voluntarily induced sets. *Psychol. Monogr.*, 1931, **42**, No. 188.
24. SEARS, R. R. Effect of optic lobe ablation on the visuo-motor behavior of goldfish. *J. comp. Psychol.*, 1934, **17**, 233–265.
25. SGONINA, K. Vergleichende Untersuchungen über den bedingten Reflex. *Zool. Anz.*, 1939, **125**, 248–251.
26. SGONINA, K. Vergleichende Untersuchungen über die Sensibilizierung und den bedingte Reflex. *Z. Tierpsychol.*, 1939, **3**, 224–247.
27. UFLAND, I. M. Die Reflexerregbarkeit des Frosches während des Umklammerungsreflexes. *Arch. ges. Physiol.*, 1929, **221**, 605–622.
28. WICKENS, D. D., & WICKENS, C. A study of conditioning in the neonate. *J. exp. Psychol.*, 1940, **26**, 94–102.

DISCRIMINATIVE CONDITIONING. I. A DISCRIMINATIVE PROPERTY OF CONDITIONED ANTICIPATION

BY WILLIAM K. ESTES

University of Minnesota

Anticipation has been defined as a reaction to a current stimulus arising from the fact that the latter has in the past been followed by a reinforcing stimulus (1). Anticipatory states may in general be of two types: (*a*) that in which the anticipated stimulus is a disturbing stimulus, such as an electric shock; and (*b*) that in which the anticipated stimulus is a positively reinforcing stimulus, such as food. In the former case, a number of properties of the anticipatory state have been experimentally established. The work of Liddell and others (3, 4) has shown that a previously neutral stimulus which repeatedly precedes a shock acquires the capacity to arouse autonomic reactions, such as changes in breathing rate or electrical resistance of the skin, as well as skeletal reflexes which originally occurred only in response to the shock itself. And an experiment recently reported by the writer in collaboration with Dr. B. F. Skinner (1) has demonstrated that anticipation of a shock produces widespread changes in strength of other responses in the organism's repertoire. That is, a stimulus which has in the past been followed by a disturbing stimulus becomes the occasion for a state of 'anxiety' which involves, not only the autonomic and specific skeletal changes traditionally associated with emotion, but also a depression in strength of other responses which constitute the organism's normal behavior in the situation, even though these responses have never themselves been associated with the disturbing stimulus.

In the second case—that in which the anticipated stimulus is positively reinforcing—much less is known about the properties of the anticipatory state. The sequence of events in which the presentation of food is repeatedly preceded by some originally neutral stimulus is, of course, characteristic of the classical Pavlovian conditioning procedure. The Pavlovian type of experiment has, however, typically yielded information concerning only an isolated local effect of the changed state of the organism which comes to be aroused by the 'conditioned stimulus,' and has never been designed to reveal any more general properties of the anticipatory state. The existence of such general properties has been indicated by an experiment recently

reported by K. C. Walker (**6**), which showed that a stimulus that has been the basis of a discrimination in the case of one response will exhibit a similar discriminative effect [1] when transferred to a new response. The question now arises whether a stimulus, S_1, which has become the occasion for a state of anticipation of a reinforcing stimulus, S_2, may not exert a discriminative effect upon any response in the organism's repertoire which has in the past been reinforced by S_2, even though no response has ever been associated with S_1 in establishing a discriminated operant. Specifically, the experiment to be described was designed to determine whether a tone which repeatedly precedes the presentation of food, where the presentation of food does not reinforce any specified response, will acquire the capacity to exert a discriminative effect upon a lever-pressing response which has previously been reinforced with food but has never been associated with the tone prior to the test trial.

Subjects and Procedure

The conditioning apparatus has been described in detail elsewhere (**2**, **5**). Each rat was enclosed during the daily one-hour experimental period in a light-proof and nearly sound-proof box containing a lever which could easily be depressed and a food magazine which could be set to reinforce responses to the lever with small (.05 gm.) pellets of food. Under the procedure of periodic reinforcement, a control clock was set to reinforce single responses to the lever every four min., intervening responses going unreinforced. A curve representing number of responses vs. time was recorded for each rat. The discriminative-tones were produced by phones attached to a 60 cycle A.C. transformer.

The subjects were 8 albino rats, about 90 days old. The experimental and control groups each included 2 males and 2 females. The animals were fed dry food for one hour immediately after running each day. Water was available in the living cages at all times.

After preliminary conditioning of the pressing response, both groups were given four experimental hours of periodic reinforcement, following which the experimental group received discriminative training for three days under the following procedure. The animals were placed in the experimental boxes at the usual time, but with the levers removed from the boxes. After a few minutes, the tones were turned on and allowed to sound for 60 sec.; then a pellet of food was delivered to each rat by means of the magazine, and the tones were turned off. This combination of tone and food was repeated ten times during each of the three discriminative training periods, and at no time during these periods were the levers present in the boxes. At each presentation, the magazine was set to deliver a food pellet after 60 sec. of tone irrespective of the momentary behavior of the rat. On the day following these training periods, both groups were returned to the boxes for one hour of experimental extinction of the lever-pressing response. The levers were again present in the boxes, and responses to the lever were recorded, but the food magazines were disconnected so that all responses went unreinforced.

The tone was presented to the rats in the experimental group for two ten-minute intervals during this extinction period in order to determine whether it would now exert any discriminative effect upon their behavior with respect to the lever. The tone was not presented to the control group because a 'normal' extinction curve, undisturbed by any complicating factor, was desired to provide an orderly baseline against which any effect of the tone upon the behavior of the experimental animals could be measured. Other experiments (**1**, **6**) furnish an adequate control against any unconditioned facilitative effect of the tone upon the pressing response. In fact, there is some indication that a tone may exert a slight depressive effect upon the rate of pressing

[1] The term 'discriminative effect' of a stimulus implies merely that the stimulus has become an occasion for the occurrence of the response. (*Cf.* **5**, pp. 177–189.)

when first introduced, an effect which it was desirable to avoid in the present experiment since it would have spuriously increased any difference between the two groups which might appear as a result of the discriminative conditioning of the experimental animals.

Results and Discussion

Mean extinction records for the two groups are shown graphically in Fig. 1. For convenience in analysis of the data, the extinction

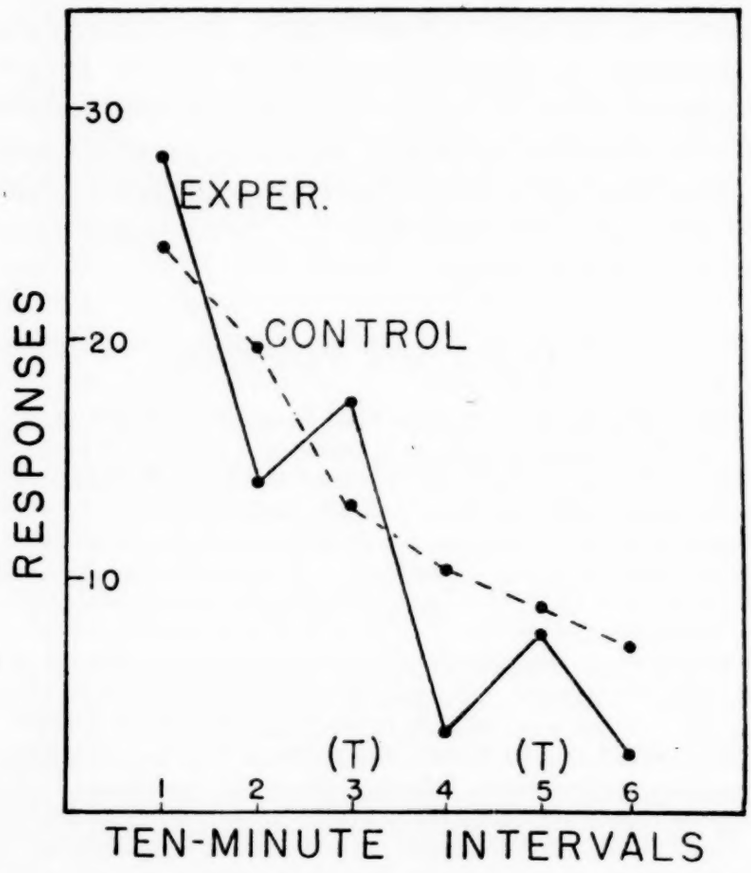


Fig. 1. Discriminative effect of the tone upon the rate of responding of the experimental group as compared to the controls. Mean extinction records for the two groups plotted by ten-minute intervals. The tone was turned on during the third and fifth intervals.

period has been divided into six ten-min. intervals. The tone was sounding during the third and fifth of these intervals for the experimental group. It will be seen from the figure that the tone exerted a marked discriminative effect upon the pressing response although it had never previously been associated with the response. When the tone was presented to the experimental rats at the beginning of the third ten-min. interval of extinction, the rate of responding showed a distinct rise, and during the subsequent silent interval an equally distinct depression. The same pattern was reproduced in the fifth and sixth intervals. This effect is similar to that of a discriminative stimulus which has previously been correlated with the reinforcement

of an operant (5). The control group, as one should expect, produced a typical negatively accelerated extinction curve, with the rate of responding declining regularly throughout the hour.

A chi-squared comparison of the distribution of responses during the portion of the extinction period subsequent to the introduction of the tone, with the mean number of responses during the intervals when the tone was sounding and those during the two subsequent silent intervals grouped for the experimental animals and compared with the corresponding values for the controls in a four-fold table, yields a value of chi-squared which would be expected less than three times in a hundred in a chance distribution.

A more detailed comparison of the two groups is summarized in Table 1. Since the absolute rates of responding during extinction are

TABLE 1

MEAN DIFFERENCES IN AMOUNTS OF RESPONDING BETWEEN SUCCESSIVE TEN-MINUTE INTERVALS

A negative difference indicates an increase in amount of responding from one interval to the next. The tone was presented to the experimental group during periods 3 and 5.

Group	Difference				
	1–2	2–3	3–4	4–5	5–6
Experimental: Mean	9.50	−8.50	14.25	−4.24	5.00
Control: Mean	4.00	7.25	2.50	1.50	1.50
t	1.30	3.84	2.24	4.08	.94
P	.12	.004	.03	.003	.19

affected by motivational differences and other variables which are not relevant to the present analysis, the effects of the tone upon the behavior of the Experimental animals are shown best by comparing the two groups with respect to changes in amount of responding from one ten-minute interval to the next, rather than with respect to amounts of responding within each interval. In Table 1 these changes have been measured by subtracting the number of responses occurring within a given interval from the number in the immediately preceding interval. The significances of the differences between corresponding values for the two groups have been evaluated by means of Student's *t*-test. According to this test, the increases in rate of responding for the experimental group as compared to the controls during the intervals when the tone was sounding are statistically highly significant, as is the depression in rate of responding for the experimental rats during the first silent period following a tone period. The question as to whether the values of *t* obtained might not be due in part to a reduction in variability on the part of the

experimental group rather than to mean differences alone may be disregarded. Irrespective of the contribution of differences in variability, the high values of t for these comparisons indicate significant effects of the experimental variable since the two groups do not differ significantly in total amounts of responding during extinction or in rates of responding during the portion of the extinction period prior to the introduction of the tone.

A comparison of the early portions of the two extinction records furnishes some information as to the effect of the discriminative training upon the reserve of the pressing response. The periodic reinforcement of the response prior to the discriminative conditioning had, of course, the usual effects of increasing the rate of pressing and of building up a considerable reserve of potential behavior which was then available for emission during the extinction period when reinforcement had been discontinued. The question arises whether the discriminative conditioning produces either an increase or a decrease in this reserve, aside from any effect upon the response in the presence of the stimulus. The fact that the two groups begin the extinction period at similar rates and do not differ significantly in total number of responses emitted during the hour suggests that there has been little effect of either sort. On the other hand, the more rapid decline in rate of pressing during the first 20 minutes of extinction for the experimental group as compared with the controls, suggests that the response might have extinguished more rapidly in the experimental animals if the tone had not been introduced during extinction. Since the difference is not statistically reliable, no definite conclusion can be drawn at present. In any case, it is clear that the primary effect of the correlation of tone and food during discriminative conditioning is to make the reserve of the pressing response, which had previously been reinforced with food, more readily available in the presence of the tone than in its absence, irrespective of the presence or absence of an effect upon the total size of the reserve.

One objection which might be raised to this interpretation of the results is that the lever-pressing response had become associated with the experimental situation prior to the discriminative training, and that even though the levers were withdrawn from the boxes during the tone-food presentations, the tendency to make the response might still have been present and therefore directly associated with the tone. An experiment designed to test this possibility will be reported shortly.

Summary

The principal conclusion to be drawn from the present study is that an originally neutral stimulus, S_1, which is repeatedly followed

by a positively reinforcing stimulus, S_2, becomes the occasion for an anticipatory state, one important property of which is a discriminative effect upon any response in the organism's repertoire which has in the past been reinforced by S_2. In the experiment described, a series of presentations of a tone followed by food to a group of rats resulted in the conditioning of an anticipatory state to the tone, the primary index being the discriminative effect of the tone upon a lever-pressing response which had previously been reinforced with food but in no way associated with the tone. During a test period subsequent to the series of tone-food combinations, the rate of lever pressing was markedly increased during intervals when the tone was sounding and depressed during silent intervals, although the response had never been associated with the tone prior to the test period.

(Manuscript received August 24, 1942)

References

1. Estes, W. K., & Skinner, B. F. Some quantitative properties of anxiety. *J. exp. Psychol.*, 1941, **29**, 390–400.
2. Heron, W. T., & Skinner, B. F. An apparatus for the study of animal behavior. *Psychol. Rec.*, 1939, **3**, 166–176.
3. Liddell, H. S. The conditioned reflex. In *Comparative psychology* (F. A. Moss, Ed.). New York: Prentice-Hall, 1934.
4. Mowrer, O. H. Preparatory set (expectancy)—some methods of measurement. *Psychol. Monogr.*, 1940, **52**. Pp. 43.
5. Skinner, B. F. *The behavior of organisms: an experimental analysis.* New York: Appleton-Century, 1938.
6. Walker, K. C. Effects of a discriminative stimulus transferred to a previously unassociated response. *J. exp. Psychol.*, 1942, **31**, 312–321.

A QUANTITATIVE AND QUALITATIVE ANALYSIS OF THE ELECTROENCEPHALOGRAMS OF STUTTERERS AND NON-STUTTERERS [1, 2]

BY HARTWELL E. SCARBROUGH [3]

Los Angeles, California

Introduction

Of the various theories of stuttering, the one advanced by Travis (**26**) has been especially provocative of research. According to this theory, there exists an insufficiently dominant lateral cortical gradient and there is a resultant rivalry of function between the two hemispheres; since single control over speech functions is not possible in this instance, the rivalry is expressed as stuttering. While the theory states that there exists within the stutterer an imbalance of functioning of the central nervous system, the imbalance is to be inferred from peripheral measures such as breathing movements (**24, 25, 30**), masseter muscle action potential (**24, 25, 27**), etc. (**15, 17**).

It must also be stated that such a theory postulates a more or less continuous imbalance, present during silence as well as during speech or attempted speech, the most severe imbalance being associated with the phenomenon known as stuttering. Normal, or non-stuttering, speech is assumed to be related to a more precise neurophysiologic adjustment, and normal speakers are hence relatively free of such imbalance.

Attempts to measure by more or less direct means the state of the central nervous system are relatively recent. Since the advent of electroencephalography, Travis and Knott (**28, 29**) have made two quantitative investigations of the electroencephalograms of stutterers and normal speakers; their studies include comparisons of speech and non-speech, as well as of stuttering and non-stuttering speech. In the main, their work has been less unequivocal than one might desire, especially in their attempts to differentiate stutterers and non-stutterers during silence. If these two behaviorally different groups are assumed to be physiologically different in any basic way, and if

[1] From the Laboratory of Physiological Psychology, Department of Psychology, University of Iowa.

[2] A condensation of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology, in the Graduate College of the State University of Iowa, August, 1939.

[3] The author wishes to express his gratitude to Dr. Wendell Johnson and Dr. John R. Knott for their guidance of this research.

the differences assumed are cortical, as alleged, some systematic differences in their electrocortical activity might be expected.

The present investigation was concerned with this type of differentiation: that which is made not on the basis of speech behavior, or of measures obtained during speech, but on the basis of the electroencephalographic behavior of the supposedly dominant hemisphere with peripheral behavior minimized. Specifically, four questions were attacked:

(1) Are there statistically significant differences in the mean number of cortical potentials per second in the left (dominant) hemisphere between stutterers and normal speakers? (Frequency measures are known to differentiate other groups of behavior deviates (**6, 8, 9, 14**).)

(2) Are there significant differences in mean variability of the number of cortical potentials per second between stutterers and non-stutterers? (If stutterers are more variable than non-stutterers, as has been suggested by Morely (**14**), on the basis of peripheral measures, then there might be such a difference in the variability of cortical activity.)

(3) If such differences do appear, are they consistent, *i.e.*, will they appear upon reëxamination?

(4) Are there significant differences with regard to the number of records in each group presenting qualitatively abnormal electroencephalographic activity of the type reported in epileptics (**8**), psychotics (**6, 9**) and behavior-problem children (**14**)? (Such differences characterize patients exhibiting episodic disorders, even though no abnormal episode is temporally correlated with the recording; since the disorder of stutterers may be described as paroxysmal and episodic, it is at least conceivable that they might be differentiated during non-episodic, *i.e.*, non-stuttering, intervals.)

Apparatus and Procedure

The technique followed in this problem was that of studying the cortical potential activity of two groups of individuals, stutterers and normal speakers. There were 20 subjects in each group. All subjects were right-handed. In the normal speaking group, there were 4 females and 16 males; the average age was 23 years, ranging from 20 to 26 years. In the stuttering group, there were 2 females and 18 males; their average age was 25 years, ranging from 19 to 34 years.

The equipment consisted of three well-matched, four-stage, resistance-capacitance-coupled amplifiers, the outputs of which were coupled to a Grass ink-writing oscillograph, which recorded at a rate of 30 mm. per sec., by means of special four-tube matching amplifiers. The time constant of each amplifier was 0.67 sec.

The subjects reclined on a cot in a semi-sound-proofed, electrostatically shielded room (which was itself in a room adjacent to the recording room). An intercommunicating system provided constant verbal contact between subject and experimenter, with selective contact between experimenter and subject.

The electrodes were flattened lead pellets held in place by collodion. Commercial electrode paste interposed between scalp and electrode surface assured good contact. The electrode

resistance was kept below 15,000 ohms at all times. All records were taken from the left hemisphere; since the subjects were all right-handed this presumably was the dominant hemisphere. The electrodes were over the left occipital area, the left motor area, and the left frontal area. The 'inactive' electrodes were placed on the lobes of the two ears.

The subjects were instructed to recline on a cot with eyes closed and to refrain from going to sleep. Any sleep-like activity occurring on the records called for immediate verbal questioning of the subject as to whether he felt drowsy or sleepy. If there was evidence that this was the case, an appropriate preceding section of the record was marked off as being out of consideration for quantitative measurement.

Two fifteen minute electroencephalographic recordings were made for each subject, the second record being run at least a week after the first.

Analysis of the Data

The data secured were analyzed in the following manner:

Quantitative measurement was made in terms of number of waves per second. Two 30-sec. samples from each record, involving the three areas being studied, were selected for measurement. The first sample was taken from the first five minutes of each record and the second sample from the last five minutes. The 30-sec. samples were marked off into sections one second long and the numbers of waves in each section computed. Any observed change of potential falling within the following limits was called a wave: if the drop in a potential was equal to one-half its rise, it was called a wave. Any blurred, blocked or unreadable material was not considered, except where the blur was easily observed to be in a wave form.

The reliability of measurement was checked by re-reading the three areas on three unmarked records (90 sec. per record) picked at random. The coefficients of correlation for the two readings of the three records were .97, .94, and .97.

The reliability of the test was measured by correlating the averages of the measures of the first sample of each run with those of the second sample, and by correlating the average of the first run with those of the second run. The inter-sample correlations were .86 for the left occipital area, .83 for the left motor area, and .88 for the left frontal area. The inter-run coefficients of correlation were .65 for the left occipital area, .63 for the left motor area and .55 for the left frontal area.

The qualitative analysis of the records was made in terms of the criteria of abnormality described by Gibbs, *et al.* (**8**).

Results

Qualitative Analysis

For all subjects the number of waves per second of cortical activity was measured for a total of 4,800 seconds. The means and standard deviations for the two 30-sec. samples of each record were computed for all three areas. Group means, or means of these individual means, were computed for each run and for the two runs combined (means of the two runs averaged together) for the three areas. These data are presented in Tables 1, 2, and 3.

TABLE 1

Mean Number of Waves Per Second for Normal Speakers and Stutterers, First Run

	L. Occipital			L. Motor			L. Frontal		
	A.M.	S.D.	*N*	A.M.	S.D.	*N*	A.M.	S.D.	*N*
Normal speakers	11.83	1.85	1200	12.15	1.96	1200	12.48	2.23	1200
Stutterers	11.88	1.44	1200	12.58	1.79	1200	12.65	2.40	1200

TABLE 2

MEAN NUMBER OF WAVES PER SECOND FOR NORMAL SPEAKERS AND STUTTERERS, SECOND RUN

	L. Occipital			L. Motor			L. Frontal		
	A.M.	S.D.	*N*	A.M.	S.D.	*N*	A.M.	S.D.	*N*
Normal speakers..	12.38	1.88	1200	12.90	2.00	1200	12.98	2.31	1200
Stutterers.....	12.85	2.01	1200	13.73	2.35	1200	13.83	2.93	1200

TABLE 3

MEAN NUMBER OF WAVES PER SECOND FOR NORMAL SPEAKERS AND STUTTERERS, COMBINED RUNS

	L. Occipital			L. Motor			L. Frontal		
	A.M.	S.D.	*N*	A.M.	S.D.	*N*	A.M.	S.D.	*N*
Normal speakers..	12.09	1.75	2400	12.55	1.81	2400	12.73	2.05	2400
Stutterers.....	12.34	1.52	2400	13.15	1.80	2400	13.22	2.34	2400

Figs. 1, 2, and 3 present the distribution curves (percent of *N* plotted against number of waves per second) for the combined runs.

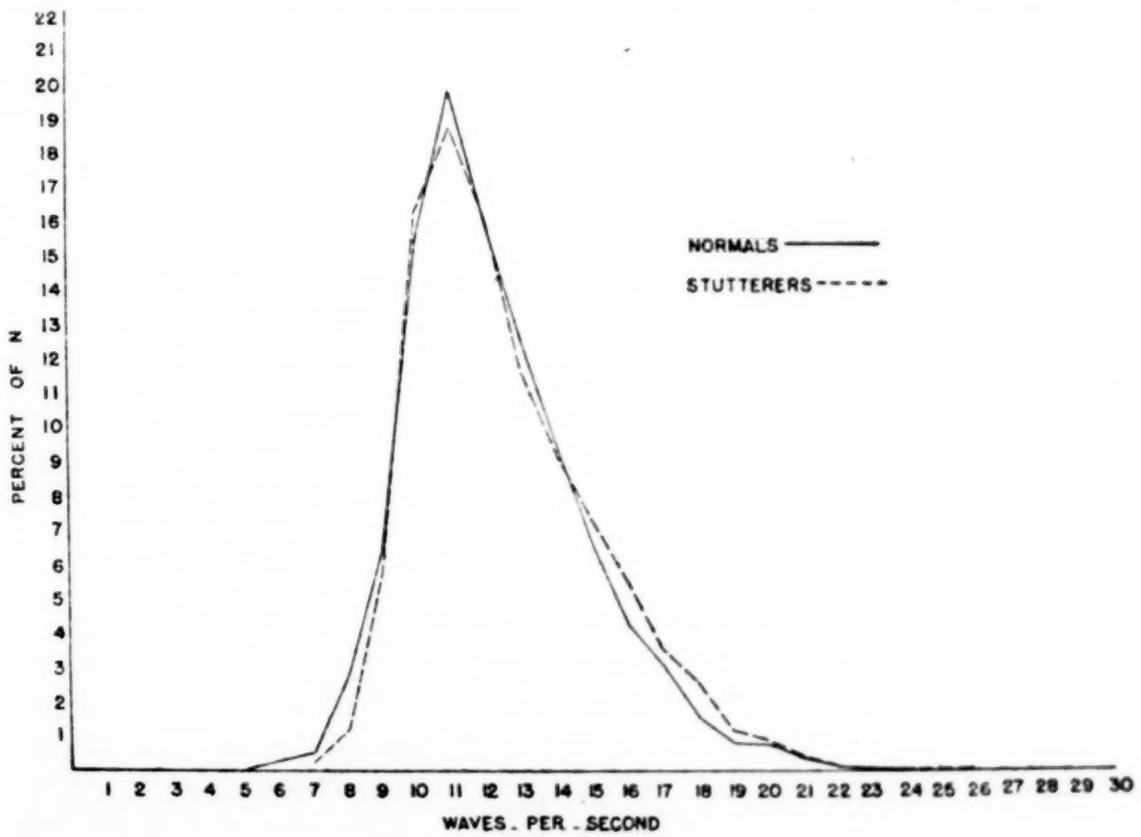


FIG. 1. Number of waves per second plotted against percentage of total number of waves (N = 2400) for the left occipital area for normal speakers and stutterers, combined runs.

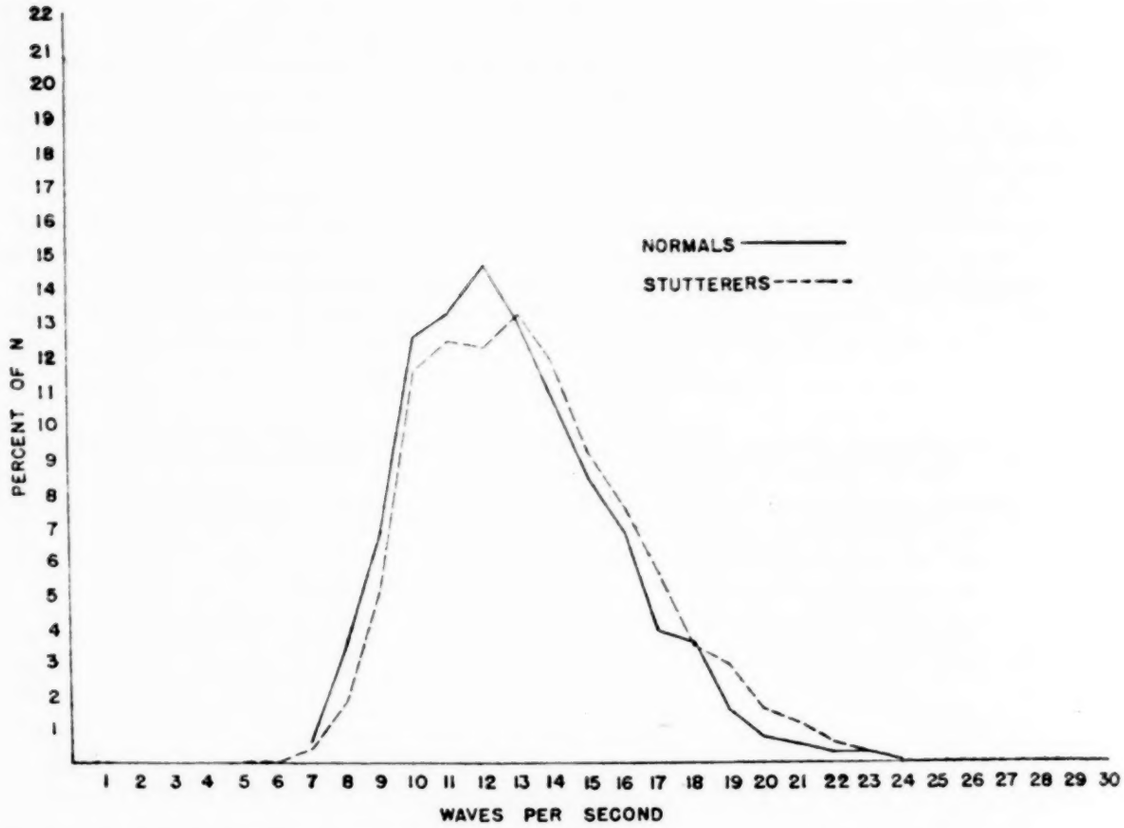


FIG. 2. Number of waves per second plotted against percentage of total number of waves (N = 2400) for the left motor area for normal speakers and stutterers, combined runs.

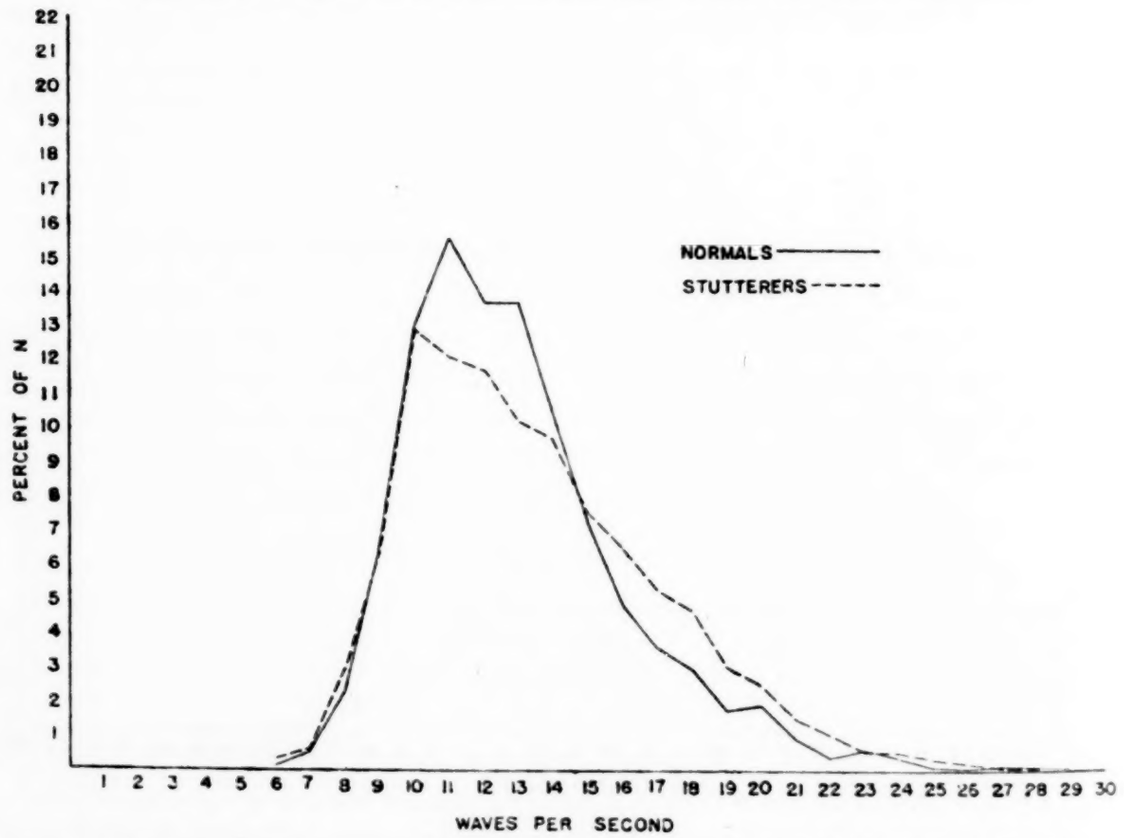


FIG. 3. Number of waves per second plotted against percentage of total number of waves (N = 2400) for the left frontal area for normal speakers and stutterers, combined runs.

It seems obvious from these figures that the curves for the two groups are not essentially different.

Since an extremely low mean number of waves per second would point to pathological activity, it is of interest to note that no means were recorded below 8 waves per second. Two subjects (one stutterer and one normal), presented mean frequency values of between 8 and 9 waves per second. (In neither of these two records were any definite qualitative abnormalities observed.)

From the data on group means in Tables 1, 2, and 3, the significance of the difference between the mean values for normal speakers and stutterers was computed. Fisher's *t* formula was used to express the significance of the difference because of the small size of the sample ($N = 20$) and the consequent need for correcting for unreliable standard deviations. Tables 4, 5, and 6 present the values of *t* and their significance.

TABLE 4

SIGNIFICANCE OF DIFFERENCES IN MEAN NUMBER OF WAVES PER SECOND, STUTTERERS VERSUS NORMAL SPEAKERS, FIRST RUN

	Fisher's *t*	*P**
Left occipital	.09	betw. .4 and .5
Left motor	.70	betw. .4 and .5
Left frontal	.23	betw. .8 and .9

TABLE 5

SIGNIFICANCE OF DIFFERENCES IN MEAN NUMBER OF WAVES PER SECOND, STUTTERERS VERSUS NORMAL SPEAKERS, SECOND RUN

	Fisher's *t*	*P*
Left occipital	.75	betw. .4 and .5
Left motor	1.17	betw. .2 and .3
Left frontal	.99	betw. .3 and .4

TABLE 6

SIGNIFICANCE OF DIFFERENCES IN MEAN NUMBER OF WAVES PER SECOND, STUTTERERS VERSUS NORMAL SPEAKERS, COMBINED RUN

	Fisher's *t*	*P*
Left occipital	.47	betw. .6 and .7
Left motor	.59	betw. .5 and .6
Left frontal	.69	betw. .4 and .5

From these results, it can be seen that the differences between the mean number of waves per second for the two groups, normals and stutterers, do not approach significance on either the first or second runs on any of the three areas.

Since the confidence to be placed in a result depends not only on the magnitude of the mean value obtained, but equally on the agree-

* *P* in this and following tables should be .05 or less before the difference can be regarded as approaching significance.

ment between parallel experiments, the differences obtained on the first and second runs between the two groups were added together and a mean taken, and Fisher's *t* formula applied to this mean difference. The results for this treatment are shown in Table 7.

TABLE 7

SIGNIFICANCE OF MEAN DIFFERENCES IN MEAN NUMBER OF WAVES PER SECOND, NORMALS VERSUS STUTTERERS, FIRST AND SECOND RUN

	Mean Difference	*t*	*P*
Left occipital	.26	1.2	betw. .4 and .5
Left motor	.63	3.1	betw. .1 and .2
Left frontal	.51	1.5	betw. .3 and .4

It will be observed that for no area was a difference approaching significance reported as a result of averaging the differences for the two runs and computing the significance of this mean difference.

Since it has been hypothesized that the stutterer is an unstable organism and since Morley (**15**) has reported the stutterer to be more variable than the non-stutterer in breathing and eye-movement behavior, it was thought to be important to ascertain whether the electroencephalograms of the two groups would significantly differentiate them with regard to their variability. The variability of the two groups was compared by computing the mean standard deviation of the two groups. Tables 8, 9, and 10 give these results.

TABLE 8

MEAN STANDARD DEVIATION OF MEAN NUMBER OF WAVES PER SECOND FOR NORMAL SPEAKERS AND STUTTERERS, FIRST RUN

	L. Occipital		L. Motor		L. Frontal	
	A.M.	*N*	A.M.	*N*	A.M.	*N*
Normal speakers	1.65	1200	2.03	1200	2.05	1200
Stutterers	1.72	1200	2.06	1200	2.22	1200

TABLE 9

MEAN STANDARD DEVIATION OF MEAN NUMBER OF WAVES PER SECOND FOR NORMAL SPEAKERS AND STUTTERERS, SECOND RUN

	L. Occipital		L. Motor		L. Frontal	
	A.M.	*N*	A.M.	*N*	A.M.	*N*
Normal speakers	1.66	1200	2.03	1200	2.09	1200
Stutterers	1.94	1200	2.06	1200	2.20	1200

TABLE 10

MEAN STANDARD DEVIATION OF MEAN NUMBER OF WAVES PER SECOND FOR NORMAL SPEAKERS AND STUTTERERS, COMBINED RUNS

	L. Occipital		L. Motor		L. Frontal	
	A.M.	*N*	A.M.	*N*	A.M.	*N*
Normal speakers	1.75	2400	2.14	2400	2.26	2400
Stutterers	2.03	2400	2.36	2400	2.59	2400

The values are seen to be in agreement for both groups, with the stutterers showing slightly more variability in each case. To ascertain whether this difference was a true difference, Fisher's *t* formula was applied to the differences between the groups for the values for each run and each area. These differences are given in Tables 11, 12, and 13.

TABLE 11

SIGNIFICANCE OF DIFFERENCES IN MEAN STANDARD DEVIATIONS OF MEAN NUMBER OF WAVES PER SECOND, STUTTERERS VS. NORMALS, FIRST RUN

	Fisher's *t*	*P*
Left occipital	.4	betw. .6 and .7
Left motor	.2	betw. .8 and .9
Left frontal	.8	betw. .4 and .5

TABLE 12

SIGNIFICANCE OF DIFFERENCES IN MEAN STANDARD DEVIATIONS OF MEAN NUMBER OF WAVES PER SECOND, STUTTERERS VS. NORMALS, SECOND RUN

	Fisher's *t*	*P*
Left occipital	1.9	betw. .05 and .1
Left motor	.4	betw. .6 and .7
Left frontal	.6	betw. .5 and .6

TABLE 13

SIGNIFICANCE OF DIFFERENCES IN MEAN STANDARD DEVIATIONS OF MEAN NUMBER OF WAVES PER SECOND, STUTTERERS VS. NORMALS, COMBINED RUNS

	Fisher's *t*	*P*
Left occipital	1.65	betw. .09 and .1
Left motor	11.47	betw. .1 and .2
Left frontal	1.43	betw. .1 and .2

It will be noted that no significant differences in the mean standard deviation are reported in the above tables for the first, second, and combined runs for any area. That is, the stutterers were not significantly more variable than the normal speakers during silence.

Again the influence of agreement between parallel experiments (first and second runs) was taken into account, as it was with the data on differences in mean values. The differences in mean standard deviation for the first run and the second run were averaged and the

t formula applied to this mean difference. The results are given in Table 14.

TABLE 14

SIGNIFICANCE OF MEAN DIFFERENCES IN MEAN STANDARD DEVIATION, NORMALS VERSUS STUTTERERS, FIRST AND SECOND RUNS

	Mean Diff.	Fisher's *t*	*P*
Left occipital	.17	1.7	betw. .3 and .4
Left motor	.04	4.0	betw. .1 and .2
Left frontal	.14	4.7	betw. .1 and .2

It will be observed that in no area did the difference in mean standard deviation approach significance as a result of this method of averaging the differences for the two runs. If these two groups are differentiated from the standpoint of intra-group variability of cortical functioning, this differentiation is not detectable during the non-speech activity of the two groups.

Since stuttering is highly intermittent and variable (from minute to minute, from day to day, and from week to week) it was argued that if the variability and intermittency was a function of an instability of cortical functioning, the electroencephalograms of stutterers might show significant differences in their degree of correlation from sample to sample and from record to record. The significance of the difference between the correlations for the first sample of each record with the second (10 minutes apart), and the first record against the second record (at least one week apart) was computed for stutterers versus normal speakers. The results are given in Tables 15 and 16.

TABLE 15

SIGNIFICANCE OF DIFFERENCES BETWEEN CORRELATIONS FOR STUTTERERS AND CORRELATIONS FOR NORMALS FOR FIRST RUN VERSUS SECOND RUN RECORDS

	Difference in Corresponding *z*	Standard Error of Difference	Ratio
Left occipital	.38	.34	1.1
Left motor	.22	.34	0.6
Left frontal	.22	.34	0.6

TABLE 16

SIGNIFICANCE OF DIFFERENCES BETWEEN CORRELATIONS FOR STUTTERERS AND CORRELATIONS FOR NORMALS FOR FIRST SAMPLE VERSUS SECOND SAMPLES

	Difference in Corresponding *z*	Standard Error of Difference	Ratio
Left occipital	.37	.34	1.1
Left motor	.00	.34	0.0
Left frontal	.12	.34	0.4

It will be noted that this comparison of the degree of correlation between first runs and second runs and between the first samples and second samples did not serve to differentiate the two groups since the differences did not approach statistical significance.

These findings indicate that cortical potential activity does not vary from sample to sample in a given run, or from run to run, in this group of stutterers significantly more than in this group of normal speakers. Variability of cortical functioning, as it was measured by the above set of measurements and within the limits of the conditions of the experiment, does not serve to differentiate stutterers from normal speakers.

Qualitative Analysis

In the group of stutterers three records were classified as 'abnormal.' In all of these, 6–8 per sec activity of the psychomotor type was observed.

In the group of normals one such record was seen.

Discussion

The present data are of interest from two points of view: the theoretical and the practical. From the theoretical standpoint, these data do not support the contention that the dominant cortical hemisphere of the stutterer functions in a manner different from that of the non-stutterer during other-than-speech activity (in the present instance, during minimal stimulation). This statement is, of course, subject to the obvious qualification that the measures utilized may have been sufficiently gross to make differentiation impossible. From the practical viewpoint, which is that of the speech clinician engaged in therapeutic labor, these data should condition remedial techniques, for unless it can be demonstrated that there are operationally definable neurophysiological differences between stutterers and non-stutterers, the speech re-education of the former group should not be encumbered by assumptions not amenable to test and the members of this group should not be told that they are pathophysiologically different from non-stutterers. The fact remains, of course, that other analytic devices may serve to differentiate these two groups electroencephalographically, and it is to be hoped that further investigation will be carried on.

Summary and Conclusions

1. A study was made to determine whether stutterers can be differentiated from normal speakers by electroencephalographic techniques. Electroencephalograms were recorded from the left occipital,

motor, and frontal areas of the brain from 20 stutterers and 20 normal speakers. Two records 15 min. long, run at least a week apart, were secured from each individual. Two 30-sec. samples were selected for quantitative measurement from each record. Measurement was made in terms of the number of waves per second. The records were also analyzed qualitatively for the presence of the abnormalities known to be related to neurophathological activity.

2. There were no statistically significant group differences in mean number of waves per second for any of the three areas. These results were the same for the first run records, the second run records, and the two runs combined.

3. A comparison of the mean standard deviation of the mean number of waves per second for the two groups showed no statistically significant differences for any of the three areas. These results were the same for the first run, the second run, and the combined runs.

4. There were no significant differences in the values of correlations obtained between the two groups for first run versus second run records, or first sample versus second sample of the same records, indicating that stutterers were not more variable from sample to sample (ten minutes apart) and from record to record (one week apart) than the normal speakers.

5. Qualitative abnormalities were frequent in neither group. Although three stutterers did show definite abnormalities, as opposed to only one normal speaker, 17 other stutterers did not. A much larger percentage difference would be necessary to make any conclusions meaningful.

(Manuscript received September 19, 1942)

References

1. Berry, M. F. Twinning in stuttering families. *Hum. Biol.*, 1937, **9**, 328–346.
2. Blackburn, W. B. A study of voluntary movements of the diaphragm, tongue, lips, and jaw, in stutterers and normal speakers. Unpublished Master's Thesis, Univ. Iowa, 1928.
3. Bryngelson, B. Psychological problems in stuttering. *Ment. Hyg., N. Y.*, 1937, 21, 631–639.
4. Child, C. M. *Physiological foundations of behavior.* New York: Henry Holt and Co., 1924.
5. Cobb, S., & Cole, E. M. Stuttering. *Physiol. Rev.*, 1939, **19**, 49–62.
6. Davis, P. A., & Davis, H. The electroencephalograms of psychotic patients. *Amer. J. Psychiat.* (In press)
7. Garvey, C. R. Conditioned respiratory changes. *Psychol. Bull.*, 1932, **29**, 676–677.
8. Gibbs, F. A., Davis, H., & Lennox, W. G. The electroencephalogram in epilepsy and in conditions of impaired consciousness. *Arch. Neurol. Psychiat., Chicago*, 1935, **34**, 1133–1140.
9. Gibbs, F. A., Gibbs, E. L., & Lennox, W. G. The likeness of cortical dysrhythmias of schizophrenia and psychomotor epilepsy. *Amer. J. Psychiat.*, 1938, **95**, 255–269.
10. Grass, A. M., & Gibbs, F. A. The Fourier transform of the electroencephalogram. *Proc. Am. physiol. Soc., 5th Annual meeting in Baltimore*, 1938, p. 80. (In press)
11. Hoagland, H. Electroencephalograms in schizophrenia. *Arch. Neurol. Psychiat., Chicago*, 1938, **39**, 210–215.

12. Hull, C. The factor of the conditioned reflex. In Murchison, C. F., *Handbook of experimental psychology*. Worcester: Clark Univ. Press, 1934, pp. 395–415.
13. Hunsley, Y. Dysintegration in the speech musculature of stutterers during the production of non-vocal temporal pattern. *Psychol. Monogr.* 1937, **49**, 32–49.
14. Jasper, H. H., Fitzpatrick, C. P., & Solomon, P. Analogies and opposites in schizophrenia and epilepsy; electroencephalographic and clinical studies. *Amer. Journ. Psychiat.* (In press)
15. Morley, A. An analysis of associative and predisposing factors in the symtomatology of stuttering. *Psychol. Monogr.*, 1937, **49**, 50–107.
16. Morris, D. W. Position as a factor of attentional clearness in relation to stuttering. *J. Speech Disorders.*
17. Moser, H. M. A qualitative analysis of eye movement during stuttering. *J. Speech Disorders*, 1938, **3**, 31–139.
18. Patzay-Lieberman, L. Bewegungsterapie bei Stottern. *Z. Kinderforsch.*, 1937, **46**, 337–345.
19. Quinan, C. Stammering and left handedness. *J. exp. Psychol.*, 1938, **23**, 304–312.
20. Raney, E. Reversed lateral dominance in identical twins. *J. exp. Psychol.*, 1938, **23**, 222–225.
21. Scott, H. D. Hypnosis and the conditioned reflex. *J. gen. Psychol.*, 1916, **23**, 89–117.
22. Seth, G. An experimental study of the control of the mechanism of speech, and in particular of that of respiration in stuttering subjects. *Brit. J. Psychol.*, 1934, **24**, 375–388.
23. Solomon, Jasper, H. H., & Bradley, E. Studies in behavior of problem children. *Proc. Boston Soc. Psychiat. Neurol.*, 1937. (In press)
24. Steer, M. D. Symptomatologies of young stutterers. *J. speech Disorders*, 1937, **1**, 3–13.
25. Strother, C. A study of the extent of dyssnergia occuring during the stuttering spasm. *Psychol. Monogr.*, 1937, **49**, 108–127.
26. Travis, L. E. *Speech pathology*. New York: D. Appleton, 1931.
27. Travis, L. E. Dissociation of homologous muscle functioning in stuttering. *Arch. Neurol. Psychiat., Chicago*, 1934, **31**, 127–133.
28. Travis, L. E., & Knott, J. R. Brain potentials from normal speakers and stutterers. *J. Psychol.*, 1936, **2**, 137–150.
29. Travis, L. E., & Knott, J. R. Bilaterally recorded brain potentials from normal speakers and stutterers. *J. speech Disorders*, 1935, **2**, 239–241.
30. Travis, V. A study of horizontal dysintegration of breathing during stuttering. *Arch. Speech.*, 1936, **1**, 158–165.
31. Van Riper, C. Effect of devices of minimizing stuttering on the creation of symptoms. *J. abnorm. (soc.) Psychol.*, 1937, **32**, 185, 192.
32. Watson, J. B. The place of the conditioned response in psychology. *Psychol. Rev.*, 1916, **23**, 89–117.
33. West, R. A. A neurological test for stutterers. *J. Neurol. Psychopath.*, 1929, **38**, 114–123.

MEDIATED GENERALIZATION AND THE INTERPRETATION OF VERBAL BEHAVIOR: II. EXPERIMENTAL STUDY OF CERTAIN HOMOPHONE AND SYNONYM GRADIENTS [1]

BY JOHN P. FOLEY, JR. AND CHARLES N. COFER

The George Washington University

In a previous paper (1), the writers presented a theoretical account of the conditioned response mechanism of mediated generalization, and suggested that many phenomena of verbal behavior might be understood in terms of mediated as well as non-mediated generalization. The present paper reports an investigation designed to determine whether some of the suggested theoretical relationships may be demonstrated experimentally.

In the above-mentioned report, it was pointed out that if generalization effects should operate as described, they might occur along any of several different dimensions. Thus, if a given word be reinforced, we should expect generalization effects to appear in its synonyms, homophones (1, Figs. 2 and 3), antonyms, and other classes of words which the subject has learned to associate with it in some way.[2] The present investigation is limited to certain synonym and homophone relationships. The study was designed specifically to discover whether generalization may be demonstrated to occur to (1) more than one homophone of a reinforced word, (2) a synonym of the reinforced word, and (3) a synonym of a synonym of the reinforced word which is not, however, itself a synonym (or other semantic relative) of the reinforced word. The first two problems involve generalization to stimuli directly related semantically or homophonically to the reinforced word, whereas the third problem involves semantic generalization to a stimulus only indirectly related to the reinforced word—two transformations removed from the original reinforced word, *cf.* (1).

The writers have previously (1) summarized the experiments directly concerned with generalization along synonym and homophone gradients. To our knowledge, all of these experiments have been limited to the demonstration of such generalization along a *single* synonym or homophone gradient from a given word, the synonym

[1] The present paper represents part of a research project being conducted by tne writers under a grant-in-aid from the American Association for the Advancement of Science.

[2] For a list of certain of these possible gradients along which mediated generalization might be expected to occur, *cf.* (1).

gradient in each case being but *one degree* of transformation or 'mediation' removed from the original, reinforced word. The present study thus represents an extension of the concept of mediated and non-mediated generalization to include generalization along more than one homophone gradient from a given word and along synonym gradients two transformations removed from the original word.[3] Attention is also called to the fact that additional data are reported in the literature which are relevant to the problem as studied by the method here employed, although no attempt will be made to summarize such material in the present report.

METHOD

Theoretically, if a given word is reinforced, *i.e.*, responded to in a particular stimulus situation, not only will its reaction potential be incremented, but the reaction potentials of all other words related to it will likewise be increased to some degree. This increased reaction potential of the other words constitutes a generalization effect, mediated and/or non-mediated, which should be experimentally demonstrable. As previously suggested (1), several methods might be employed to study this phenomenon,[4] only one of which has been used in the present study.

1. *Materials.*—A list of 10 unrelated words (presentation or P-list) was assembled such that from each word two homophones and one synonym could be derived. In addition, each of the synonyms had a familiar meaning other than that of the original word, and thus a synonym of each synonym was available, although the second synonym was not a synonym of the word in the original list. In each case, care was taken to minimize other complicating semantic relationships.[5] According to the hypothesis advanced by the writers, the repetition (reinforcement) of these derived or related words should increase the reaction potentials of the words in the original list; if this should occur, one should expect the latter to be more easily learned after the presentation of a related list than after the presentation of an unrelated (control) list. Thus, to test generalization along the above-mentioned gradients, four related and one unrelated reinforcement lists were constructed as follows: (1) the H_1–R list was composed of 10 words each one of which was a homophone of one word in the original (P) list; (2) the H_2–R list was composed of 10 words each one of which was a homophone of one word in the original (P) list (there were no words common to the H_1–R and H_2–R lists); (3) the S_1–R list was composed of 10 words each one of which was a synonym of one word in the original (P) list; (4) the $S_2(2)$ list was composed of 10 words each one of which was a synonym of a word in the S_1–R list but not a synonym of any word in the original (P) list; (5) the C–R (control) list was composed of 10 words unrelated to the words in the P-list in any evident or formal way, directly or indirectly. These 5 lists, along with the P-list, may be found under the appropriate headings in Table 1.

In addition to the above lists two others were utilized. One consisted of 8 numbers (spelled out) which served as a 'buffer' or practice list. The other was a list of ten Christian names which was used to equate the subjects in different groups for 'learning ability.' The buffer and equation lists may be found in Table 1 under the headings 'B.' and 'E.L.,' respectively. It might be added that numbers and names were used for these purposes so as to minimize the possibility of generalization effects from these sources.

[3] Likewise, it would be theoretically important to investigate mediated generalization along more than one synonym gradient from the same presentation word, as well as along alternate or interlocking synonym and homophone gradients.

[4] Four methods for studying the operation of mediated generalization in verbal behavior were outlined in the previous paper (1). In the present experiment the first of these methods was employed.

[5] The writers recognize the possible existence of generalization gradients arising from the instructions as well as the possible existence of intra-serial, 'circular' gradients. Examples of the latter are *fruit-plant* (S_1–R list) and *palm-reed* (C–R list).

TABLE 1

LISTS OF WORDS USED IN THE EXPERIMENT

The word order (top to bottom) shown is that of the first presentation of each list throughout the experiment. Only the R-lists were changed in order at any time, as described in the text. Numbers in parentheses indicate the direct or indirect relationships between words in the R-lists and words in the P-list.

B.	E.L.	P.	Reinforcement Lists				
			H_1–R	H_2–R	S_1–R	$S_s(2)$–R	C–R
eleven..	edith	(1) sent	(3) pare	(3) pair	(3) fruit	(3) result	palm
six.....	bernard	(2) vein	(4) sew	(4) so	(4) plant	(4) factory	set
forty...	louise	(3) pear	(7) write	(7) rite	(7) just	(7) barely	reed
three...	vivian	(4) sōw	(2) vain	(2) vane	(2) vessel	(2) ship	very
sixteen..	joseph	(5) reign	(8) noes	(8) nose	(8) apprehends	(8) arrests	numb
nine....	dorothy	(6) sees	(9) meat	(9) mete	(9) join	(9) enlist	me
twelve..	henry	(7) right	(10) dō	(10) doe	(10) batter	(10) bruise	day
five....	calvin	(8) knows	(6) seas	(6) seize	(6) looks	(6) appearance	snap
	julia	(9) meet	(5) rain	(5) rein	(5) rule	(5) principle	rope
	samuel	(10) dough	(1) cent	(1) scent	(1) dispatched	(1) killed	spire

Mean no. of letters per word.....	5.0	5.9	4.2	3.8	3.8	5.9	6.7	3.7
Range in no. of letters per word....	3–7	5–7	3–5	2–5	2–5	4–10	4–10	2–5

2. *Procedure.*—The buffer, equation, and P-list were presented once to each of the 125 subjects. The same intra-serial order (that of Table 1) was used for every subject in the case of these lists. Five sub-groups of 25 subjects each were used for the presentation of the reinforcement lists. One group received four presentations of the H_1–R list, another group four presentations of the H_2–R list, and so on. The subjects were told that they would not be asked to remember the reinforcement lists. During the four presentations of a given reinforcement list, the order of the words was changed each time. The same order was used for each presentation for every subject in a given group. Furthermore, the position throughout the reinforcement presentations of all words having a relationship (direct or indirect) to a given word in the P-list was the same for all reinforcement lists. The serial order for the first R-list presentations is shown in Table 1 (reading from top to bottom). Thus, if the fifth word in one of these reinforcement list (*e.g.*, H_1–R) occurred as the first word on the second presentation of that reinforcement list, all other fifth words in the other reinforcement lists (including C–R) would likewise occur as the first words in the second presentations of the other reinforcement lists, and so on.

The regular procedure was as follows: The buffer list was presented, and *S* wrote his recall. The E.L. list was then presented at once, and the *S*'s written recall of that list was immediately scored. As it was desirable to keep the learning ability of the five sub-groups equivalent, the determination of which reinforcement list to present was made on the basis of the E.L. score. (It will be seen from Table 2 that the five sub-groups were practically identical in respect to mean E.L. score.) Then a given R-list was presented four times, and this was followed immediately by the single presentation of the P-list; *S* then wrote the words which he recalled from the P-list, after which verbal reports were secured. During the presentation of all lists, *S* pronounced each word aloud as it appeared in the aperture of the memory drum.

The experiment was conducted in an inside, windowless room, with low noise level. The lists were typed on white paper and presented one word at a time by means of the Gerbrands memory drum (3, No. M1–4). By using a multiple exposure shield, previously described (2), it was necessary to change the paper only once—between the presentation of the E.L. list and the presentation of the first R-list. The drum was set at the one and one-half second exposure speed for the presentation of all words. The entire experiment required approximately 15–18 minutes for each individual *S*.

The specific *instructions* given to each subject were as follows. At the beginning of the experiment, *E* said: "Watch this opening (indicating it) in which a series of words will be exposed

TABLE 2

DATA CONCERNING THE EQUIVALENCE IN 'LEARNING ABILITY' OF THE 5 SUB-GROUPS WITH RESPECT TO EQUATION LIST SCORE

Reinforcement List	N	'Equation List' Performance			
		No. Correct	No. Incorrect	Range in No. Correct	Mean No. Correct
H_1–R	25	164	5	3–10	6.56
H_2–R	25	165	9	5–8	6.60
S_1–R	25	166	3	4–8	6.64
$S_a(2)$–R	25	165	3	4–9	6.60
C–R	25	165	6	5–9	6.60

one at a time. As each word is exposed, say it aloud and think of its meaning. When all the words have been exposed, you will be asked to put down on paper as many of the words of this series as you can remember. Are there any questions?" After the presentation of the first (buffer) list, *E* said: "Now, put down as many words of this series as you can remember; their order is not important." The instructions for the equation list were the same as for the buffer list. Before the presentation of the reinforcement lists, *E* said: "Now *this* time say each word aloud as you did before and think of its meaning, but you will *not* be called upon to remember them. If the letter *o* with a long line over it occurs in a word, this *o* is to be pronounced as a long *o*, as in *mōde*." No other comments were made during the four presentations of the reinforcement lists, other than to repeat the above instructions before each presentation. Before the presentation of the P-list, *E* said: "Now *this* time say each word aloud and think of its meaning; when all the words have been exposed, you will be asked to put down as many of the words in *this series* as you can remember." After the P-list had been presented, *E* said, "Now put down as many of the words of this series as you can remember."

Following the *S*'s recall of the P-list, he was asked the following questions in order: 1. "Do you have any comments?" 2. "Did you make use of any aids or other devices to help your recall? What devices?" 3. "Did you notice any relationships between words on any of the lists? What relationships?"

3. *Scoring.*—In the case of both the E.L. and the P-list, the score was computed in terms of number of words correctly recalled, as in the method of retained members. The order in which the words were recalled was not considered. All legible words were scored, and slight mistakes in spelling did not count as errors so long as it was clearly evident that the correct word and not some other word (especially the homophone) was intended.

4. *Subjects.*—125 *S*s were used in the experiment, 56 male and 69 female. They were obtained from psychology classes but were naïve with respect to the purpose of the experiment. As previously described, the five sub-groups were equated for 'learning ability' on the basis of scores on the Equation List (*cf.* Table 2). They were also approximately equated as to age, sex ratio, number of courses in psychology, and number of years in college.

RESULTS

The major results of the experiment, in terms of the number of words correctly and incorrectly recalled by each of the five equated sub-groups, are presented in Table 3. The reliabilities of the differences between the mean or median of each of the four experimental sub-groups and that of the control sub-group are found in Table 4.[6]

[6] Since normality of distribution could not be assumed, $\sigma_{Mdn.}$ was computed throughout by the following formula:

$$\sigma_{Mdn.} = \frac{i\sqrt{N}}{2F}.$$

TABLE 3

ANALYSIS OF PERFORMANCE OF THE 5 SUB-GROUPS IN THE RECALL OF THE PRESENTATION LIST

Reinforcement List	N	No. Correct	No. Incorrect	Range Correct	Mean Correct	S.D. Correct	Median Correct	Q Correct
H_1–R	25	173	13	4–10	6.72	1.61	6.67	1.24
H_2–R	25	141	30	3–8	5.64	1.49	5.42	1.28
S_1–R	25	147	14	3–9	5.88	1.46	5.92	1.11
$S_a(2)$–R	25	131	13	3–7	5.24	1.14	5.42	0.90
C–R	25	120	26	1–7	4.80	1.55	4.45	1.15

TABLE 4

RELIABILITIES OF DIFFERENCES IN CENTRAL TENDENCY BETWEEN EACH OF THE 4 EXPERIMENTAL SUB-GROUPS AND THE CONTROL SUB-GROUP IN THE RECALL OF THE PRESENTATION LIST

Reinforcement List	Analysis in Terms of Mean					Analysis in Terms of Median				
	σM.	Difference from C–R Mean	σdiff.	Significance Ratio	Probability Value	σMdn.	Difference from C–R Mdn.	σdiff.	Significance Ratio	Probability Value
H_1–R	.33	1.92	.80	2.39	99.16	.83	2.22	.87	2.55	99.46
H_2–R	.30	0.84	.79	1.07	85.77	.42	0.97	.49	1.98	97.61
S_1–R	.30	1.08	.78	1.38	91.62	.42	1.47	.49	3.00	99.87
$S_a(2)$–R	.23	0.44	.74	0.59	72.24	.42	0.97	.49	1.98	97.61
C–R	.32	—	—	—	—	.25	—	—	—	—

It will be noted that computations of central tendency, variability, and reliability are reported in terms of both mean and median measures. Although there is a general similarity in results obtained with these two measures, certain differences are readily apparent. Owing to the small number of cases (25) in each sub-group and the somewhat irregular distributions, the median is a more representative measure of central tendency. The present discussion, therefore, will be based upon an analysis in terms of the median, although a corresponding analysis in terms of the means can be found in Tables 3 and 4.

The median number of words correctly recalled by the control sub-group, *i.e.*, the sub-group receiving the unrelated reinforcement list, is 4.45; this may be used as a basis for evaluating the scores of the other, experimental sub-groups. It will be noted that the two sub-groups receiving reinforcement with homophone lists correctly recalled a median of 6.67 and 5.42 words, respectively. The former median is 2.22 words larger than that of the control, and the sigma of the difference is .87, with a significance ratio of 2.55; the probability-value of this difference is 99.46, indicating the chances in 100 that such a difference could not have arisen as a sampling error from a true difference of zero. The significance ratio in the case of the other

homophone-reinforced sub-group, whose median is 0.97 words larger than that of the control, is 1.98, corresponding to a probability-value of 97.61. The sub-group receiving reinforcement from synonyms of the presentation words shows a median of 5.92 words correctly recalled, the difference of 1.47 from the control median yielding a significance ratio of 3.00 and probability-value of 99.87. Finally, the sub-group reinforced with synonyms two transformations removed from the original presentation list is found, in terms of the median, to recall 0.97 more words than the control sub-group, with resulting significance ratio of 1.98 and probability-value of 97.61.

On the basis of the experimental method and treatment of data here described, it is apparent that the probability that such generalization effects might have arisen from chance factors is very small. Not only is the median number of words correctly recalled significantly greater in the case of the experimental sub-groups, but certain additional results in the case of the particular experimental sub-groups conform rather closely to theoretical expectation. Chief among these considerations is the fact that generalization to the first-order synonyms is stronger than to the second-order synonyms. No summary conclusion can be drawn, however, regarding the relative strength of generalization to homophones and synonyms. The median number of words correctly recalled for the first homphone list is greater than that for the one list of first-order synonyms, although the greater variability of the former makes its significance ratio smaller than that of the synonym list. Conflicting conclusions, furthermore, are obtained if the significance ratios of the differences between the medians are compared with those based upon the means. It is nevertheless clear that the median or mean number of words correctly recalled is significantly greater in the case of both homophone and synonym sub-groups, although it is obvious that the particular words employed as homophones or synonyms will undoubtedly influence the amount of generalization to a marked extent. In this connection, attention is called to the difference between generalization effects in the case of the two homophone lists. This difference appears to indicate that, contrary to Razran's interpretation (*cf.* 4), homophones do not represent simple instances of physical or non-mediated generalization, but rather represent such generalization complicated by other (including semantic) factors. It will also be noted that the presentation words used in the present study were selected so that two homophones of each word were available; if this narrow restriction had not been imposed, generalization along a single homophone gradient as well as along the synonym gradients would possibly have been even more readily demonstrable (owing to the more obvious homophone and synonym relationships).

The writers have tabulated the number of correct recalls of each presentation word for each of the five sub-groups. The relative frequency of recall of a given presentation word did not differ markedly among the sub-groups. Differences in the preëxperimental reaction potentials of the words, as well as differences in serial position, make further interpretation of this question unwarranted.

An analysis of the individual incorrect recalls in each of the sub-groups in many cases reveals generalization effects theoretically similar to those previously described as producing the correct recalls. Examples of the possible operation of these generalization gradients are cited below. Of the 13 incorrect recalls in the H_1–R sub-group (*cf.* Table 3), 10 were repetitions of words in the R-list, and 3 might be explained in terms of homophone and/or quasi-homophone gradients. Of the 30 incorrect recalls in the H_2–R sub-group, 18 were repetitions of words in the R-list, and the remainder seem explicable on the basis of homophone gradients (one also involving a synonym relationship). In the S_1–R sub-group, 4 of the 14 incorrect recalls were repetitions of words in the R-list; the remaining 10 could possibly be attributed to the operation of homophone, quasi-homophone, synonym, and other gradients such as antonym, genus-species, object-attribute, verb-object, and words in the same category (*e.g.*, *vessel-pier*). Only one of the 13 incorrect recalls in the $S_s(2)$ sub-group was a repetition of a word in the R-list; 8 of the remaining 12 errors could possibly be explained on the basis of other gradients, as above, and in the other 4 cases the relationship was obscure. Of the 26 incorrect recalls in the C–R sub-group, 16 were repetitions of the R-list, 7 might be accounted for in terms of the above-mentioned gradients, and 3 showed no apparent relationship.

Typical examples of incorrect recalls theoretically attributable to the operation of such generalization gradients were the following: (*a*) incorrect recall of *so*, a homophone of *sew* in R-list or of *sōw* in P-list (found in H_1–R sub-group); (*b*) incorrect recall of *row*, a quasi-homophone of *sōw* in P-list (found in C–R sub-group); (*c*) incorrect recall of *beat*, a synonym of *bruise* in R-list (found in $S_s(2)$ sub-group); (*d*) incorrect recall of *weed*, a species of *plant* in R-list (found in S_1–R sub-group); (*e*) incorrect recall of *apples*, in same class (fruit) as *pear* in R-list (found in $S_s(2)$ sub-group); (*f*) incorrect recall of *reap*, a word in 'rural category' with, and a possible antonym of *sōw* in P-list (found in $S_s(2)$ sub-group); (*g*) incorrect recall of *seeds*, a possible synonym of, or an object of, *sōw* in P-list (found in C–R sub-group); (*h*) incorrect recall of *dear*, a homophone of *deer* (not presented) whose synonym is *doe* in R-list (found in H_2–R sub-group).

Finally, let us consider the evidence contained in the *S*s' reports. It will be recalled that three questions were asked of each *S* after the

experiment was concluded. The purpose of these questions was to discover if the *S* had noticed the relationship between the R-list and the P-list. In the case of the H_1–R and H_2–R sub-groups, the relationship was definitely observed in most cases. In the former group, 24 *Ss* reported it correctly, and one apparently saw the relationship in a few of the words. Of the *Ss* in the H_2–R sub-group, 18 reported the relationship clearly, 4 reported seeing it in some words, and 3 gave no evidence of having discerned it. On the other hand, the *Ss* in the S_1–R and $S_s(2)$–R sub-groups seldom described the relationship. Only 2 of the S_1–R *Ss* saw it clearly, 6 saw it in a few words, and the rest gave no evidence of having discovered it. One member of this group spontaneously made the following comment 24 hours after he had participated in the experiment: "There was an unconscious (*sic*) connection between those words, like *batter-dough*. I didn't realize it until today." None of the *Ss* in the $S_s(2)$–R sub-group gave any evidence of having seen the relationship. In the C–R sub-group, of course, no relation was present, and in only 2 cases did the *Ss* attempt to construct a relationship; these were based on quasi-homophonimity. It should be noted that even in the cases in which the *Ss* were unable to report recognition of the actual relationship, there was definite evidence of the generalization effects from the reinforcement lists to the presentation list.

Numerous other comments were made by the *Ss* in answer to the questions, such as speculation concerning the nature of the experiment and postulation of incorrect relationships. It seems unnecessary to list or to attempt to interpret these comments.

Summary

This experiment was designed to demonstrate generalization (*a*) along more than one homophone gradient from a given stimulus word, (*b*) along a synonym gradient from the stimulus word, and (*c*) along a synonym gradient two degrees of transformation removed from the original stimulus word. Results seem to indicate generalization in all of these dimensions, and to this extent confirm the writers' previous theoretical analysis (1).

(Manuscript received September 24, 1942)

References

1. Cofer, C. N., & Foley, J. P., Jr. Mediated generalization and the interpretation of verbal behavior: I. Prolegomena. *Psychol. Rev.*, 1942, **49**, 513–540.
2. Foley, J. P., Jr., & Cofer, C. N. An inexpensive multiple-exposure extension for the simple memory drum. *J. Exp. Psychol.*, 1942, **31**, 438–439.
3. Gerbrands, R. *Scientific instruments* (mimeographed).
4. Razran, G. H. S. A quantitative study of meaning by a conditioned salivary technique (semantic conditioning). *Science*, 1939, **90**, 89–90.

LEVEL OF MASTERY AND REMINISCENCE IN PURSUIT LEARNING[1, 2]

BY CLAUDE E. BUXTON

Northwestern University

Experimental Design

This study was patterned after the studies of reminiscence in verbal learning published by Ward (**5**), Hovland (**3**), Melton and Stone (**4**), and others. These investigators practiced both their experimental and control groups to a chosen criterion of mastery. The recall trial for the control group followed immediately after the trial on which the criterion was reached, with no special rest, but the recall trial for the experimental group was delayed any desired length of time. Since the two groups reached the same level of mastery initially, any superiority of the experimental group in recall was due, presumably, to the rest they had. It must be assumed or shown that rest as such, rather than further covert practice, is responsible for the gain.

Ward has emphasized the fact that successive practices by an individual characteristically show an irregular increase in mastery, during the over-all improvement. Consequently, if all *S*s cease practice when they reach some mastery criterion arbitrarily designated by *E*, many will reach it at a temporary peak of performance. This means that in the experimental design used by Ward and others it is very likely that the control group will show something of a drop on its recall trial. Presumably the same factor operates in the experimental group but is over-balanced by their gain.

Procedure

This experiment employed a modification of the Koerth pursuit rotor, with *S* learning to use a stylus in following a target moving in a circular path. A Cenco counter recorded the amount of time actually spent in contact with the target. Each trial was 15 sec. in length, followed by a 30 sec. rest. 'Volunteers' from psychology classes made up each of four groups of *S*s, the selection being randomized by simply assigning each *S* to a group before he came into the experimental room. He was instructed in holding the stylus and maintaining suitable posture, and then asked to continue to practice until *E* decided that he had practiced long enough. He was informed of his score after each trial. *S*s in group A ceased practice after they secured a score of 90 or better—5 percent of the possible score—on any trial. *S*s in groups B and C were treated similarly, except that the criteria for initial mastery were 360 (20 percent) and 630

[1] The substance of this paper was presented at the St. Louis meetings of the M.P.A., May, 1942.

[2] Mr. William Giles and Mr. Gordon Eckstrand assisted in this study.

(35 percent). Each of these groups then had a 10½ min. rest interval, during which they read popular magazines. They returned to practice until they met their criterion a second time. The fourth group of *Ss* served as controls for each of the first three. They practiced to the highest criterion without rest—other than the 30 sec. between trials—and, still without special rest, re-learned to it. It thus was possible from the records to find the trials on which they met each criterion, and to ascertain their scores on these and the immediately following trials (which served as recall trials after a 'no-rest' interval). It is possible that the reading did not fully prevent covert rehearsal by the experimental groups, but it was felt that such thinking as they might do about the pursuit task would not produce very great transfer after the rest interval. This possibility will be checked in later work.

Results

Reminiscence.—Fig. 1 illustrates a finding such as Ward emphasized for verbal learning. Both group B and the control group at that

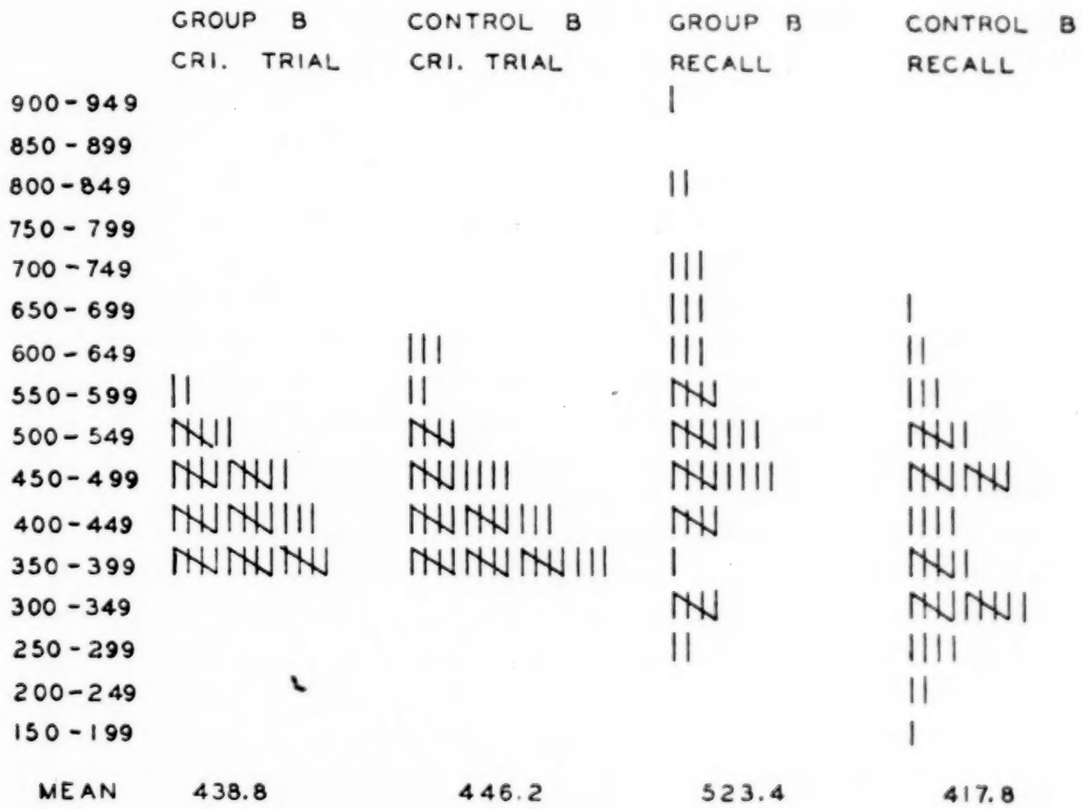


Fig. 1. Distributions of rotor scores on criterion and recall trials, for groups B and control B (criterion, 360).

level show a markedly skewed distribution on the criterion trial. On the next trial, after 30 sec. for the controls and 10½ min. for the experimental group, both groups show a more nearly normal distribution. The mean of the controls is lower than on the criterion trial, but the mean of the experimental group is clearly higher on its recall trial. (Comparisons at the A and C levels show the same type of result.)

The major results of the experiment are summarized in Table I. In the column headed 'criterion trial' is shown the matching of the

TABLE I

MEAN SCORES ON CRITERION AND RECALL TRIALS, IN EXPERIMENTAL AND CONTROL GROUPS

Group		N	Cri. Trial	S.E.m.	Recall	S.E.m.	C.R.	Chances in 100
	A	40	151.0	8.9	241.7	12.8		
Control	A	50	144.5	10.2	171.3	16.2	3.42	99+
	B	47	438.8	8.5	523.4	21.0		
Control	B	50	446.2	10.7	417.8	16.5	3.96	99+
	C	39	713.2	13.6	714.6	28.0		
Control	C	50	717.4	11.6	637.9	22.7	2.13	96+

experimental groups with the control group at the corresponding level. The criterion score for both A groups was 90, which they exceeded by about 60–70 percent, on the average. In view of the large S.E.$_{m.}$'s, the matching seems adequate. The same statement may be made concerning the matching at the B and C levels.

In the column headed 'recall' we may see the effect of the rest interval. The three control group means are lower than the experimental means, the differences being reliable for the two lower levels of learning and nearly so for the third. In accordance with the experimental design employed, these superiorities of the experimental groups are interpreted as evidence of reminiscence in their acquisition of the pursuit skill. Two remarks about these findings are pertinent: First, so far as I can determine, no study of verbal learning has produced reliable and indisputable evidence of reminiscence at an interval longer than about 5 min., if that long. The interval here was 10 min., which leads to the notion that retention curves for short intervals do not have the same form in motor learning that they do in verbal learning. Further experimentation is likely to show that the optimum interval for reminiscence in motor learning is longer than the 2 min. optimum in verbal learning. Be that as it may, there is a second pertinent remark: So far as the absolute amount of reminiscence found in this study is concerned (as shown by the size of the difference between experimental and control groups on the recall trial), it is greatest at the middle level of initial mastery employed. One can speculate that at the lowest level the Ss have not practiced enough to profit from a rest, while at the highest criterion they are beginning to approach their ceiling and therefore *cannot* improve a great deal during rest. At the middle level of mastery neither of these statements is true.

One may also compute the relative amount of reminiscence in the three experimental groups. One way of doing it is to divide the absolute amount of reminiscence (difference in mean recall for experi-

mental and control groups) by the criterion trial mid-mean of any chosen pair of groups. According to this computation, relative reminiscence decreases from 47 percent in the group with the lowest criterion to 24 percent for the middle level of mastery, and to 11 percent for the highest level of mastery before rest.

The idea will readily occur that the recall advantage of the experimental groups in this study is due to the fact that they had an opportunity to recover from fatigue. Two arguments against this may be presented: First, the individual trials were but 15 sec. long, and the rests between them were twice that length. This plan was adopted to *reduce* fatigue. Secondly, the correlation between number of trials to the designated criterion and amount of gain during the rest interval was computed for the three experimental groups. Following the view of Doré and Hilgard (1, 2), if fatigue were a factor, it is plausible that the Ss who took more trials to the criterion should be more fatigued, and should show the greatest gains during the rest. The reverse seems to be true: *eta* (used instead of *r*) for group A was .39, for group B, .39, and for group C, .53, all the relationships being negative in direction. To state the finding positively: the Ss who progressed by leaps and bounds and who reached the criterion early in the practice session tended to be the ones who showed the greatest gains during the rest interval. Fatigue thus does not seem to be an adequate explanatory device, nor are other theories of reminiscence as yet easily applied to motor learning, as McClelland [3] has shown. One rather plausible idea concerning pursuit learning is that the rest interval permits overly tense Ss to relax, thereby stepping up accuracy on the recall trial. This coördination, after all, is a rather fine one, and great muscular tension is probably a handicap.

Re-learning.—It has often been found that even though recall scores do not indicate the advantage of a rest interval, re-learning scores might do so. Table II, when compared with Table I, shows that the reverse is true in this study. The control group is well matched with the experimental groups at each of the three levels of pre-rest mastery (see the column headed 'trials to criterion'), and the rest interval produces relatively quicker re-learning in the experimental groups. Furthermore, the advantage is greatest at the middle level of mastery which was employed here. On the other hand, the effect of the rest-interval, as measured by the statistical reliability of advantages produced by it, is less when measured by re-learning scores than when measured by recall scores.

This finding, as well as those concerning reminiscence, can be said to apply only to the relatively low levels of mastery used in this study. There is no assurance that the same trends would hold if, *e.g.*, 70 or

[3] From unpublished seminar notes, very kindly loaned to me by Dr. D. C. McClelland.

TABLE II

MEAN NUMBERS OF TRIALS TO THE CRITERION AND TO RE-LEARN, IN EXPERIMENTAL AND CONTROL GROUPS *

Group		N	Trials to Cri.	S.E.m.	Trials to Relearn	S.E.m.	C.R.	Chances in 100
	A	40	4.4	.6	1.1	.05		
Control	A	50	3.5	.4	1.2	.1	.91	63+
	B	47	10.9	1.0	1.2	.1		
Control	B	50	10.4	.8	1.6	.1	2.84	99+
	C	39	16.4	1.1	1.3	.1		
Control	C	50	16.4	1.1	2.1	.3	2.53	98+

* Both trials to the criterion and for re-learning were computed by including the criterion trials themselves.

80 percent mastery before rest were required. But it does seem to be true that in pursuit-rotor learning, at least, the advantage of rest tends to disappear after the first post-rest trial.

SUMMARY

In pursuit learning, the absolute amount of reminiscence increases, then decreases, as level of mastery before a constant rest-interval rises; the relative amount of reminiscence (according to the measures employed), on the other hand, decreases in a negatively accelerated function as the initial mastery becomes greater. Reminiscence is here shown over a 10-min. interval, longer than the longest interval permitting indisputable reminiscence in verbal learning. Fatigue is not thought to be adequate as an explanation of the present results. The re-learning data substantiate the recall findings, but the rest interval was not so clearly beneficial when the re-learning index is employed.

(Manuscript received August 24, 1942)

REFERENCES

1. DORÉ, L. R., & HILGARD, E. R. Spaced practice and the maturation hypothesis. *J. Psychol.* 1937, **4**, 245–259.
2. DORÉ, L. R., & HILGARD, E. R. Spaced practice as a test of Snoddy's two processes in mental growth. *J. exp. Psychol.*, 1938, **23**, 359–374.
3. HOVLAND, C. I. Experimental studies in rote-learning theory. I. Reminiscence following learning by massed and by distributed practice. *J. exp. Psychol.*, 1938, **22**, 201–224.
4. MELTON, A. W., & STONE, G. R. The retention of serial lists of adjectives over short time-intervals with varying rates of presentation. *J. exp. Psychol.*, 1942, **30**, 295–310.
5. WARD, L. B. Reminiscence and rote learning. *Psychol. Monogr.*, 1937, **49**, No. 4.

A MERCURY SWITCH FILM TIMER

BY GEORGE E. PRICE

University of Washington

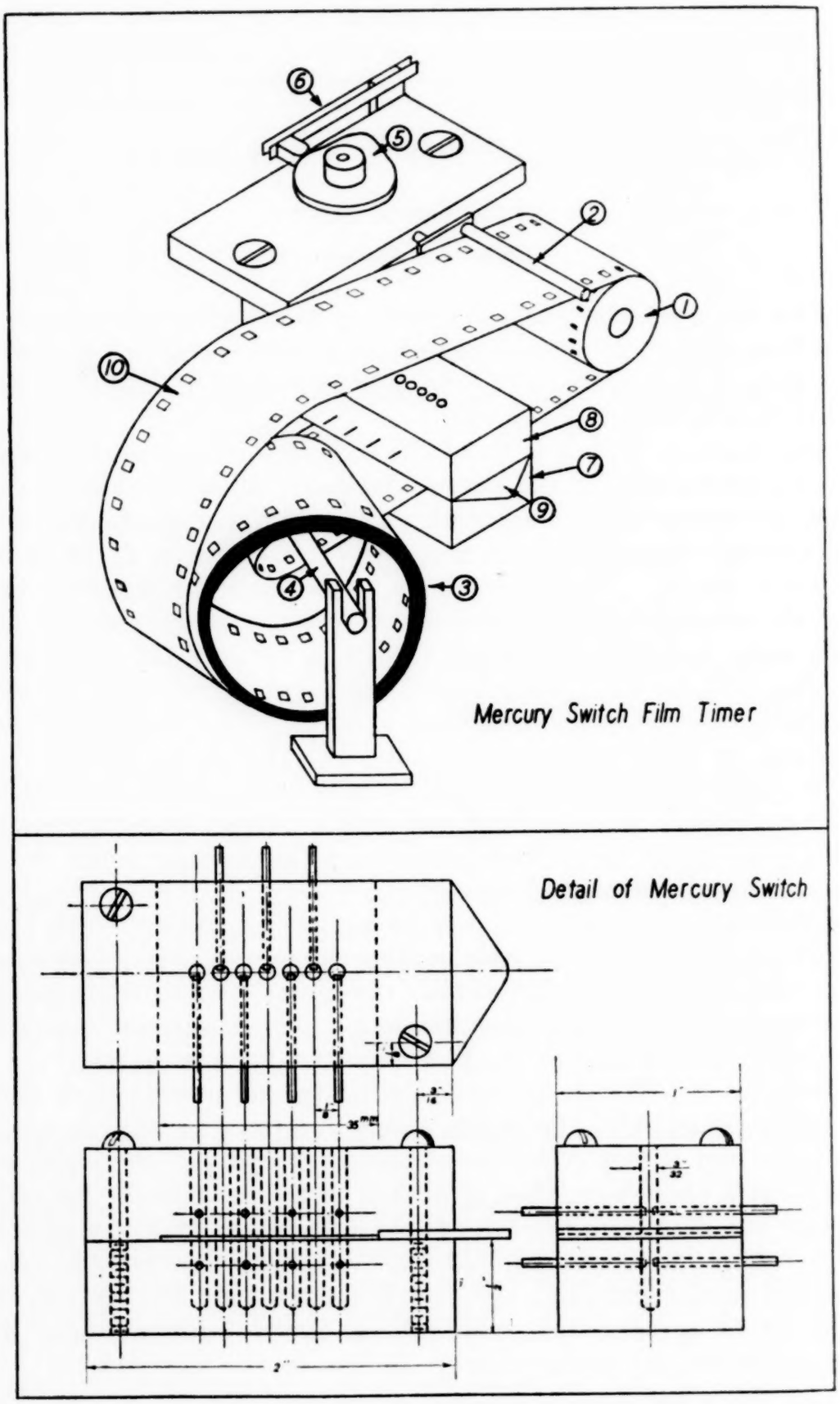
This timer enables one to control up to seven circuits through a very long series. It consists mainly of a mercury switch controlled by a strip of 35 mm film with holes punched through it, and a master switch to keep the time constant.

The mercury switch (7) and (8) consists of two bakelite blocks bolted together with a slot between them to allow the film to pass. There are seven 3/16″ holes running in a line perpendicular to the film through the upper block and about two-thirds of the way into the lower block. These holes have small wires running into them from the sides of the block, both in the lower and upper block. The wires make contact with mercury which is put into the seven holes. The film (10), in which are punched holes corresponding to the series sequence, is inserted by removing tab (9). The mercury is kept separated in the upper and lower half of the switch by the film. As a hole in the film comes in line with a hole in the switch the mercury in the upper half drops through and makes contact with the mercury in the lower half. The clearance for the film must be small so that no mercury will be carried out.

The master switch (6) is in series with the common to the lower half of the switch. By suitable gearing the master switch is synchronized with the film sprocket so that the master switch is made after the mercury switch and broken before it. This maintains a constant time interval even though the film may be a little irregular.

Holes in the film are punched with a special punch which corresponds to the switch. To enable a long film to be handled in a small space the film is first coiled into a two-inch coil and the inside end pulled out a short way, then cemented to the outer end. In action the film is pulled from the roll (3) through switch (7–8) by sprocket (1) and from there pushed back to the outside of the roll. As is indicated in the drawing, the film is guided only by the mercury switch (7–8), sprocket (1), guide roll (2), and roll support (4). The small guide roll (2) serves to keep the film on the sprocket.

This timer has been in use over a year and has offered very little trouble. It will carry about 3 amps. per circuit. The main difficulty


6
5
2
1
10
8
7
9
4
3
Mercury Switch Film Timer
Detail of Mercury Switch

experienced has been in getting the joint in the film thin enough to go through the mercury switch. This was done by cementing the film with Glacial Acetic Acid and dressing it down with a sharp knife or sandpaper.

(Manuscript received September 14, 1942)